GenCore version 4.5

TM nucleic - nucleic search, using sw model

Run on: February 18, 2001, 08:51:31 ; Search time 128.56 Seconds
(without alignments)
8643.515 Million cell updates/sec

Title: US-09-434-382-3
Perfect score: 2958
Sequence: 1 cgcgggcgtaggtgaccggc.....aataaagattgagtttgcaa 2958

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 480022 seqs, 187831343 residues

Word size : 0

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : N_Geneseq_36:*
1: /cgn1_8/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /cgn1_8/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /cgn1_8/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /cgn1_8/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /cgn1_8/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /cgn1_8/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /cgn1_8/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /cgn1_8/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /cgn1_8/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /cgn1_8/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /cgn1_8/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /cgn1_8/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /cgn1_8/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /cgn1_8/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /cgn1_8/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /cgn1_8/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /cgn1_8/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /cgn1_8/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /cgn1_8/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /cgn1_8/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /cgn1_8/gcgdata/geneseq/geneseqn/NA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	237	8.0	238	21	Z80231	Human colon cancer
	2	53	1.8	72	16	T25953	Human gene signatu
c	3	39	1.3	68	19	X11554	Human biallelic po
c	4	19	0.6	526	20	V88751	EST clone HK189.
	5	19	0.6	7642	20	V82020	Moraxella catarrha
	6	18	0.6	45	20	Z33905	Human PRO274 hybri
c	7	18	0.6	177	20	X18066	Coding sequence fo
c	8	18	0.6	531	18	T67773	H. pylori cytoplas
c	9	18	0.6	531	18	T77453	H. pylori cytoplas
	10	18	0.6	726	21	Z80749	Human colon cancer
	11	18	0.6	1040	21	Z48812	Soybean inositol 1
c	12	18	0.6	1593	18	T67992	H. pylori cytoplas
c	13	18	0.6	1626	20	X08683	Novel nucleotide s
	14	18	0.6	2186	19	V17351	Coding sequence fo
c	15	18	0.6	2450	20	X20537	Polynucleotide seq
c	16	18	0.6	2605	19	V04699	Homo sapiens 20q13
	17	18	0.6	2945	20	Z33895	Human PRO274 nucle
c	18	18	0.6	4120	20	Z09473	Human RNA helicase
	19	18	0.6	6139	19	V70354	Coding strand of n
c	20	18	0.6	49999	20	Z23903	Human LOBO homolog
	21	17	0.6	21	15	Q54684	C kappa exon prime
c	22	17	0.6	256	21	A32079	Plant microsatelli
c	23	17	0.6	300	21	A32050	Plant microsatelli
	24	17	0.6	300	21	A01382	Human colon cancer
c	25	17	0.6	304	17	T07238	Immunogen DNA from
c	26	17	0.6	330	21	Z42643	Human 5' EST isola
c	27	17	0.6	332	20	X24993	Murine Bcl-2 inter
	28	17	0.6	347	16	T24629	Human gene signatu
c	29	17	0.6	369	20	X41005	Human secreted pro
c	30	17	0.6	422	20	X24994	Murine Bcl-2 inter
c	31	17	0.6	426	21	A32069	Plant microsatelli
c	32	17	0.6	431	21	A32052	Plant microsatelli
c	33	17	0.6	559	19	V29360	Calcium ion channe
c	34	17	0.6	590	20	X24995	Murine Bcl-2 inter
c	35	17	0.6	605	20	V87198	EST clone BN180.
c	36	17	0.6	791	20	X24979	Clone GJ156 encodi
	37	17	0.6	855	21	Z91823	Streptococcus pneu
	38	17	0.6	1022	20	X40199	MAGE-4 encoding ge
	39	17	0.6	1084	15	Q67866	H6/MAGE-1 expressi
	40	17	0.6	1084	20	Z08442	H6/MAGE-1 expressi
	41	17	0.6	1094	15	Q67865	H6/MAGE-1 expressi
	42	17	0.6	1094	20	Z08441	H6/MAGE-1 expressi
	43	17	0.6	1140	20	X27353	Human secreted pro
	44	17	0.6	1173	13	Q29634	Hepatitis C virus
	45	17	0.6	1173	13	Q29635	Hepatitis C virus
	46	17	0.6	1173	14	Q43895	NANB hepatitis vir
	47	17	0.6	1209	21	Z46089	cDNA encoding a fo
	48	17	0.6	1338	20	X87593	CLYTA-MAGE-1-His f
	49	17	0.6	1341	20	X87591	Lipoprotein D-MAGE
c	50	17	0.6	1345	12	Q14450	16S RNA from ATCC
c	51	17	0.6	1557	17	T27644	cDNA encoding prot
c	52	17	0.6	1591	19	T98595	DNA encoding GTP-b
	53	17	0.6	1592	19	V42963	Streptococcus pneu
	54	17	0.6	1632	19	V68059	Neurodegenerative
	55	17	0.6	1664	20	Z25023	Murine D6 encoding
	56	17	0.6	1668	20	V72116	Mouse FAST-1 codin
	57	17	0.6	1691	20	V69719	Tumour rejection a
	58	17	0.6	1702	20	Z52876	Human prostate tum
	59	17	0.6	1724	19	V38385	Beta(1 -> 4)-N-ace
	60	17	0.6	1816	20	X85940	DNA encoding human
	61	17	0.6	1818	20	X01577	Human HCDR-1 codin
c	62	17	0.6	1906	21	A15550	Human TRAF four as
c	63	17	0.6	2160	16	T09084	Nitrite reductase
c	64	17	0.6	2284	21	Z50582	Human epidermal pr
	65	17	0.6	2418	20	X84103	E antigen precurso
	66	17	0.6	2419	13	Q32351	Antigen E gene. H
	67	17	0.6	2419	15	Q72476	Tumour rejection a
	68	17	0.6	2419	16	T05086	MZ2-MEL antigen E
	69	17	0.6	2419	20	X84112	Antigen E coding s
	70	17	0.6	2420	15	Q72472	Tumour rejection a
	71	17	0.6	2420	16	Q85435	Human melanoma ant
	72	17	0.6	2503	19	V59595	Human secreted pro
	73	17	0.6	2646	20	Z42096	Human endometrium
	74	17	0.6	2711	19	V68056	Neurodegenerative
c	75	17	0.6	2971	9	N81166	fdhF gene. synthe
	76	17	0.6	3000	12	Q13115	Encodes partial mu
	77	17	0.6	3156	19	V18471	T-cell surface ant
	78	17	0.6	3156	20	Z27969	Human CD97 protein
c	79	17	0.6	3457	14	Q48468	Nitrate reductase
	80	17	0.6	3796	18	T93499	Xenopus frog prote
	81	17	0.6	4216	8	N70558	Sequence of nitrog
	82	17	0.6	4258	18	T45351	Human colon carcin
	83	17	0.6	4308	21	Z51806	Full length expand
	84	17	0.6	4488	14	Q51426	Human FACC cDNA cl
	85	17	0.6	4567	20	V33945	Fanconi anaemia co

86	17	0.6	5674	13	Q32352	MAGE-1 nucleic aci
87	17	0.6	5674	15	Q72477	Tumour rejection a
88	17	0.6	5674	20	X84113	MAGE-1 gene. Homo
89	17	0.6	5720	17	T42117	Lats gene encoding
90	17	0.6	5720	21	Z51508	Drosophila melanog
91	17	0.6	5724	16	Q98902	Tumour rejection a
92	17	0.6	6464	14	Q48772	Arylamine N-acetyl
93	17	0.6	10240	19	V52165	Streptococcus pneu
94	17	0.6	10461	20	X20553	Polynucleotide seq
95	17	0.6	10813	18	V74675	Staphylococcus aur
96	17	0.6	11236	15	Q70447	Rat nestin gene -
c 97	17	0.6	21170	20	X20535	Polynucleotide seq
98	17	0.6	29604	18	X83005	Partial mouse WRN
c 99	17	0.6	34094	20	Z30163	Complete nucleotid
c 100	17	0.6	38734	20	Z32020	Human METH1 relate

ALIGNMENTS

RESULT 1
Z80231
ID Z80231 standard; cDNA; 238 BP.
XX
AC Z80231;
XX
DT 07-APR-2000 (first entry)
XX
DE Human colon cancer cell line SW480 cDNA clone SEQ ID NO:315.
XX
KW Human; gene expression product; diagnosis; tumour; colon cancer;
KW colorectal adenocarcinoma; cell line SW480; cell proliferation;
KW cytostatic; sarcoma; breast cancer; neoplasia; dysplasia;
KW hyperplasia; ds.
XX
OS Homo sapiens.
XX
PN WO9964576-A2.
XX
PD 16-DEC-1999.
XX
PF 09-JUN-1999; 99WO-IB01062.
XX
PR 10-JUN-1998; 98US-0088801.
XX
PA (FARB) BAYER CORP.
XX
PI Endege WO, Steinmann KE, Astle JH, Burgess CC, Bushnell SE;
PI Carroll E, Catino TJ, Derti A, Ford DM, Lewis ME, Monahan JE;
PI Schlegel R;
XX
DR WPI; 2000-087220/07.
XX
PT Novel nucleic acids, used to develop products for the diagnosis and
PT treatment of disorders involving unwanted cell proliferation,
PT particularly cancers, especially colon cancer -
XX
PS Claim 15; Page 258; 469pp; English.
XX
CC Z79917 to Z80766 represent double stranded cDNA clones isolated from the
CC human colorectal adenocarcinoma (colon cancer) cell line SW480. The
CC cDNA clones can be used to generate antisense oligonucleotides which
CC can be used for antisense therapy. Methods and products from the present
CC invention can be used for identifying and/or classifying cancerous cells
CC present in a human tumour, particularly in solid tumours, e.g. carcinomas
CC and sarcomas, e.g. breast or colon cancers. The cDNA clones can be used
CC for developing agents for the diagnosis and treatment of disorders
CC involving unwanted cell proliferation, such as neoplasia, dysplasia or
CC hyperplasia.
XX
SQ Sequence 238 BP; 55 A; 57 C; 69 G; 57 T; 0 other;

Query Match 8.0%; Score 237; DB 21; Length 238;
Best Local Similarity 100.0%; Pred. No. 4.1e-106;
Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    226 acctgcaggtggtggcagcgggtagccgggactcgggcgccgcgctctacgtcttctccg 285
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 acctgcaggtggtggcagcgggtagccgggactcgggcgccgcgctctacgtcttctccg 60

Qy    286 agttcaaccggtatctcttcaactgtggagaaggcgttcagagactcatgcaggagcaca 345
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 agttcaaccggtatctcttcaactgtggagaaggcgttcagagactcatgcaggagcaca 120

Qy    346 agttaaaggttgctcgcctggacaacatattcctgacacgaatgcactggtctaatgttg 405
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 agttaaaggttgctcgcctggacaacatattcctgacacgaatgcactggtctaatgttg 180

Qy    406 ggggcttaagtggaatgattcttactttaaaggaaaccgggcttccaaagtgtgtac 462
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ggggcttaagtggaatgattcttactttaaaggaaaccgggcttccaaagtgtgtac 237
```

RESULT 2
T25953
ID T25953 standard; cDNA to mRNA; 72 BP.
XX
AC T25953;
XX
DT 28-OCT-1996 (first entry)
XX
DE Human gene signature HUMGS08188.
XX
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
XX
OS Homo sapiens.
XX
PN WO9514772-A1.
XX
PD 01-JUN-1995.
XX
PF 11-NOV-1994; 94WO-JP01916.
XX
PR 12-NOV-1993; 93JP-0355504.
XX
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
XX
PI Matsubara K, Okubo K;
XX
DR WPI; 1995-206931/27.
XX
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
XX
PS Claim 1; Page 1967; 2245pp; Japanese.
XX
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for

```
CC   recognising different cell types.
XX
SQ   Sequence 72 BP; 25 A; 10 C; 16 G; 20 T; 1 other;


  Query Match             1.8%;  Score 53;  DB 16;  Length 72;
  Best Local Similarity   100.0%;  Pred. No. 4.4e-16;
  Matches   53;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2906 tccgagacttaacgaaaatagtatttcagctgcaataaagattgagtttgcaa 2958
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      19 tccgagacttaacgaaaatagtatttcagctgcaataaagattgagtttgcaa 71


RESULT   3
X11554/c
ID   X11554 standard; DNA; 68 BP.
XX
AC   X11554;
XX
DT   30-MAR-1999  (first entry)
XX
DE   Human biallelic polymorphic DNA fragment ESTC169.
XX
KW   Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW   detection; phenotypic typing; characteristic; infection; hereditary;
KW   autoimmune disease; cancer; inflammation; drug; therapy; medicament;
KW   treatment; marker; ss.
XX
OS   Homo sapiens.
XX
PN   WO9820165-A2.
XX
PD   14-MAY-1998.
XX
PF   05-NOV-1997;   97WO-US20313.
XX
PR   06-NOV-1996;   96US-0030455.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Hudson T,  Lander ES,  Wang D;
XX
DR   WPI; 1998-286974/25.
XX
PT   New isolated nucleic acid segments from the human genome - used for
PT   determining polymorphic forms for use in e.g. forensics, paternity
PT   testing or phenotypic typing for disease
XX
PS   Claim 1; Page 172; 310pp; English.
XX
CC   X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC   markers which have been isolated using the primers represented in
CC   X09121-X10268. The base occupying the polymorphic site is indicated by
CC   the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC   methods for determining polymorphic forms in an individual for use in
CC   e.g. forensics, paternity testing or for phenotypic typing for diseases
CC   such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC   muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC   hypercholesterolemia, polycystic kidney disease, hereditary
CC   spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC   haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC   syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC   autoimmune diseases, inflammation, cancer, diseases of the nervous
CC   system, infection by pathogenic microorganisms, and characteristics such
CC   as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC   endurance, fertility, and susceptibility or receptivity to particular
CC   drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC   segments can also be used to produce medicaments for the treatment or
CC   prophylaxis of such diseases.
XX
SQ   Sequence 68 BP; 12 A; 13 C; 21 G; 21 T; 1 other;


  Query Match             1.3%;  Score 39;  DB 19;  Length 68;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-09;
  Matches   39;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2812 tggaaacagacggcggcacctttcctctaatccagcaaa 2850
           |||||||||||||||||||||||||||||||||||||||
Db      62 TGGAAACAGACGGCGGCACCTTTCCTCTAATCCAGCAAA 24


RESULT   4
V88751/c
ID   V88751 standard; cDNA; 526 BP.
XX
AC   V88751;
XX
DT   12-FEB-1999  (first entry)
XX
DE   EST clone HK189.
XX
KW   Expressed sequence tag; secreted protein; haematopoiesis regulator;
KW   tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
KW   chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
KW   receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX
OS   Homo sapiens.
XX
PN   WO9845437-A2.
XX
PD   15-OCT-1998.
XX
PF   10-APR-1998;   98WO-US06956.
XX
PR   10-APR-1997;   97US-0837312.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Agostino MJ,  Jacobs K,  Lavallie ER,  McCoy JM,  Merberg D;
PI   Racie LA,  Spaulding V,  Treacy M;
XX
DR   WPI; 1999-070078/06.
XX
PT   New polynucleotides encoding human secreted proteins - derived from
PT   e.g. human blood, kidney, foetal lung, placenta, testes, brain,
PT   ovary, pituitary, retina and colon cDNA libraries
XX
PS   Claim 1; Page 500; 641pp; English.
XX
CC   The present sequence represents an expressed sequence tag (EST), and is
CC   a polynucleotide of the invention. The polynucleotides of the invention
CC   are all secreted EST sequences isolated from a variety of human tissue
CC   sources. The EST sequences and proteins encoded by them are predicted to
CC   have useful biological activities which would make them suitable for
CC   treating, preventing or ameliorating medical conditions in humans and
CC   animals, although no supporting data is given. Suggested activities
CC   include nutritional activity, immune stimulating or suppressing activity,
CC   haematopoiesis regulating activity, tissue growth activity,
CC   activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC   activity, cadherin/tumour invasion suppressor activity, tumour inhibition
CC   activity. The EST sequences are also stated to be useful for gene
CC   therapy.
XX
SQ   Sequence 526 BP; 127 A; 138 C; 136 G; 125 T; 0 other;


  Query Match             0.6%;  Score 19;  DB 20;  Length 526;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     742 gaagaggggtcagggactc 760
           |||||||||||||||||||
Db     272 GAAGAGGGGTCAGGGACTC 254
```

```
RESULT   5
V82020
ID   V82020 standard; DNA; 7642 BP.
XX
AC   V82020;
XX
DT   21-JUN-1999  (first entry)
XX
DE   Moraxella catarrhalis Q8 lfr region.
XX
KW   Lactoferrin receptor; lactoferrin binding protein; LBP1; LBP2;
KW   lbpA gene; lbpB gene; ORF3; infection; otitis media; sinusitis;
KW   conjunctivitis; pneumonia; bronchitis; tracheitis; emphysema;
KW   diagnosis; therapy; vaccine; Branhamella catarrhalis; ss.
XX
OS   Moraxella catarrhalis.
XX
FH   Key             Location/Qualifiers
FT   -35_signal      90..95
FT                   /*tag= a
FT   -10_signal      108..113
FT                   /*tag= b
FT   RBS             138..147
FT                   /*tag= c
FT   CDS             152..2837
FT                   /*tag= d
FT                   /gene= lbpB
FT                   /product= LBP2
FT                   /transl_except= (pos:509..512, aa:Asp)
FT                   /note= "this codon has an apparent 1 codon
FT                         insertion, which alters the reading frame"
FT                   /note= "lbpB is specifically claimed in Claim 8;
FT                         encodes W89417"
FT   -35_signal      2974..2979
FT                   /*tag= e
FT   -10_signal      2991..2996
FT                   /*tag= f
FT   RBS             3007..3016
FT                   /*tag= g
FT   CDS             3022..6024
FT                   /*tag= h
FT                   /gene= lbpA
FT                   /product= LBP1
FT                   /note= "specifically claimed in Claim 8; encodes
FT                         W89418"
FT   CDS             3067..6024
FT                   /*tag= i
FT                   /gene= lbpA
FT                   /product= LBP1
FT                   /note= "specifically claimed in claim 8; encodes
FT                         W89419"
FT   CDS             6026..7642
FT                   /*tag= j
FT                   /gene= orf3
FT                   /note= "specifically claimed in Claim 7; encodes
FT                         W89419"
XX
PN   WO9855606-A2.
XX
PD   10-DEC-1998.
XX
PF   02-JUN-1998;   98WO-CA00544.
XX
PR   08-MAY-1998;   98US-0074658.
PR   03-JUN-1997;   97US-0867941.
XX
PA   (CONN-) CONNAUGHT LAB LTD.
XX
PI   Du R, Klein MH, Loosmore SM, Wang Q, Yang Y;
XX
DR   WPI; 1999-070266/06.
DR   P-PSDB; W89417, W89418, W89419, W89420.
XX
PT   Lactoferrin receptor genes from Moraxella, especially M. catarrhalis
PT   - useful to diagnose Moraxella infection e.g. to detect otitis media
PT   due to M. catarrhalis infection and to immunise against such
PT   infections
XX
PS   Claim 8; Fig 4A-4P'; 202pp; English.
XX
CC   This polynucleotide comprises the lactoferrin receptor (lfr) locus
CC   of Moraxella catarrhalis (Branhamella catarrhalis) Q8.  There are
CC   3 tandem genes in locus, identified as lbpA, lbpB (alternative
CC   start codons) and orf3A, respectively encoding lactoferrin binding
CC   protein 2 (Lbp2, see W89417), lactoferrin binding protein 1 (Lbp1,
CC   see W89418 and W89419) and open reading frame protein 3 (ORF3, see
CC   W89420).  The lfr locus was identified following generation of a M.
CC   catarrhalis strain Q8 genomic DNA library and screening with
CC   specific hybridisation probes.  The genes and DNA sequences of the
CC   lfr locus are useful for diagnosis, immunisation, and the
CC   generation of diagnostic and immunological reagents.  Immunogenic
CC   compositions, including vaccines, based upon expressed recombinant
CC   Lbp1 and/or Lbp2 and/or ORF3, portions of these or their analogues,
CC   can be prepared for prevention of diseases caused by Moraxella.  M.
CC   catarrhalis is a causative agent of otitis media and has been
CC   associated with sinusitis, conjunctivitis and inflammatory diseases
CC   of the lower respiratory tract, such as pneumonia, chronic
CC   bronchitis, tracheitis and emphysema.
XX
SQ   Sequence 7642 BP; 2417 A; 1726 C; 1631 G; 1868 T; 0 other;

  Query Match              0.6%;  Score 19;  DB 20;  Length 7642;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     811 tcttggtgctcaaagcaaa 829
           ||||||||||||||||||||
Db    3657 tcttggtgctcaaagcaaa 3675


RESULT   6
Z33905
ID   Z33905 standard; DNA; 45 BP.
XX
AC   Z33905;
XX
DT   07-DEC-1999  (first entry)
XX
DE   Human PRO274 hybridisation probe.
XX
KW   Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
KW   probe; blood coagulation disorder; cancer; cellular adhesion disorder;
KW   secreted protein; transmembrane protein; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9946281-A2.
XX
PD   16-SEP-1999.
XX
PF   08-MAR-1999;   99WO-US05028.
XX
PR   10-MAR-1998;   98US-0077450.
PR   11-MAR-1998;   98US-0077632.
PR   11-MAR-1998;   98US-0077641.
PR   11-MAR-1998;   98US-0077649.
PR   12-MAR-1998;   98US-0077791.
PR   13-MAR-1998;   98US-0078004.
PR   17-MAR-1998;   98US-0040220.
PR   20-MAR-1998;   98US-0078886.
PR   20-MAR-1998;   98US-0078910.
PR   20-MAR-1998;   98US-0078936.
```

PR 20-MAR-1998; 98US-0078939.
PR 25-MAR-1998; 98US-0079294.
PR 26-MAR-1998; 98US-0079656.
PR 27-MAR-1998; 98US-0079663.
PR 27-MAR-1998; 98US-0079664.
PR 27-MAR-1998; 98US-0079689.
PR 27-MAR-1998; 98US-0079728.
PR 27-MAR-1998; 98US-0079786.
PR 30-MAR-1998; 98US-0079920.
PR 30-MAR-1998; 98US-0079923.
PR 31-MAR-1998; 98US-0080105.
PR 31-MAR-1998; 98US-0080107.
PR 31-MAR-1998; 98US-0080165.
PR 31-MAR-1998; 98US-0080194.
PR 01-APR-1998; 98US-0080327.
PR 01-APR-1998; 98US-0080328.
PR 01-APR-1998; 98US-0080333.
PR 01-APR-1998; 98US-0080334.
PR 08-APR-1998; 98US-0081049.
PR 08-APR-1998; 98US-0081070.
PR 08-APR-1998; 98US-0081071.
PR 09-APR-1998; 98US-0081195.
PR 09-APR-1998; 98US-0081203.
PR 09-APR-1998; 98US-0081229.
PR 15-APR-1998; 98US-0081817.
PR 15-APR-1998; 98US-0081838.
PR 15-APR-1998; 98US-0081952.
PR 15-APR-1998; 98US-0081955.
PR 21-APR-1998; 98US-0082568.
PR 21-APR-1998; 98US-0082569.
PR 22-APR-1998; 98US-0082700.
PR 22-APR-1998; 98US-0082704.
PR 22-APR-1998; 98US-0082804.
PR 23-APR-1998; 98US-0082767.
PR 23-APR-1998; 98US-0082796.
PR 27-APR-1998; 98US-0083336.
PR 28-APR-1998; 98US-0083322.
PR 29-APR-1998; 98US-0083392.
PR 29-APR-1998; 98US-0083495.
PR 29-APR-1998; 98US-0083496.
PR 29-APR-1998; 98US-0083499.
PR 29-APR-1998; 98US-0083500.
PR 29-APR-1998; 98US-0083545.
PR 29-APR-1998; 98US-0083554.
PR 29-APR-1998; 98US-0083558.
PR 29-APR-1998; 98US-0083559.
PR 30-APR-1998; 98US-0083742.
PR 05-MAY-1998; 98US-0084366.
PR 06-MAY-1998; 98US-0084414.
PR 06-MAY-1998; 98US-0084441.
PR 07-MAY-1998; 98US-0084598.
PR 07-MAY-1998; 98US-0084600.
PR 07-MAY-1998; 98US-0084627.
PR 07-MAY-1998; 98US-0084637.
PR 07-MAY-1998; 98US-0084639.
PR 07-MAY-1998; 98US-0084640.
PR 07-MAY-1998; 98US-0084643.
PR 13-MAY-1998; 98US-0085323.
PR 13-MAY-1998; 98US-0085338.
PR 13-MAY-1998; 98US-0085339.
PR 15-MAY-1998; 98US-0085573.
PR 15-MAY-1998; 98US-0085579.
PR 15-MAY-1998; 98US-0085580.
PR 15-MAY-1998; 98US-0085582.
PR 15-MAY-1998; 98US-0085689.
PR 15-MAY-1998; 98US-0085697.
PR 15-MAY-1998; 98US-0085700.
PR 15-MAY-1998; 98US-0085704.
PR 18-MAY-1998; 98US-0086023.
PR 22-MAY-1998; 98US-0086392.
PR 22-MAY-1998; 98US-0086414.
PR 22-MAY-1998; 98US-0086430.
PR 22-MAY-1998; 98US-0086486.
PR 28-MAY-1998; 98US-0087098.
PR 28-MAY-1998; 98US-0087106.
PR 28-MAY-1998; 98US-0087208.
PR 30-JUL-1998; 98US-0094651.
PR 11-SEP-1998; 98US-0100038.
XX
PA (GETH) GENENTECH INC.
XX
PI Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
XX
DR WPI; 1999-551358/46.
XX
PT New secreted and transmembrane polypeptides and their polynucleotides,
PT useful for treating blood coagulation disorders, cancers and cellular
PT adhesion disorders -
XX
PS Example 4; Page 184; 530pp; English.
XX
CC The present invention describes secreted and transmembrane polypeptides
CC and their polynucleotides. The nucleotide sequences are useful as
CC sources of probes, primers, for chromosome mapping, and for generation
CC of antisense sequences. They can also be used to create transgenic
CC animals. The proteins can be used to treat a variety of diseases and
CC disorders, depending on their function. Diseases that may be treated
CC include blood coagulation disorders, cancers and cellular adhesion
CC disorders. They may also be used to raise antibodies. Z33891 to
CC Z34338, and Y41685 to Y41774 represent polynucleotide and polypeptide
CC sequence given in the exemplification of the present invention.
XX
SQ Sequence 45 BP; 8 A; 10 C; 19 G; 8 T; 0 other;

Query Match 0.6%; Score 18; DB 20; Length 45;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 872 cattgctgctgtcaagga 889
||||||||||||||||||
Db 26 cattgctgctgtcaagga 43

RESULT 7
X18066/c
ID X18066 standard; DNA; 177 BP.
XX
AC X18066;
XX
DT 04-MAY-1999 (first entry)
XX
DE Coding sequence for human SI binding protein SNi45.
XX
KW Gastro-intestinal transport receptor; binding protein; hSI; HPT1;
KW D2H; hPEPT1; human; GI tract receptor; sucrose-isomaltase complex;
KW intestinal peptide-associated transporter; hypertension; diabetes;
KW osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
KW therapeutic agent delivery; therapy; ss.
XX
OS Homo sapiens.
XX
PN WO9851325-A2.
XX
PD 19-NOV-1998.
XX
PF 15-MAY-1998; 98WO-US10088.
XX
PR 15-MAY-1997; 97US-0046595.
XX
PA (CYTO-) CYTOGEN CORP.
PA (ELAN-) ELAN CORP PLC.
XX
PI Alvarez VL, Belinka BA, Cagney GM, Carter JM, Lambkin IJ;
PI Omahony DJ, Patterson CA, Singleton J;
XX

```
RN   WPI; 1999-009568/01.
XX
PT   New proteins that bind specifically to receptors in the
PT   gastro-intestinal tract and related nucleic acid - chimaeras and
PT   antibodies, used to deliver therapeutic or diagnostic agents to, or
PT   through, the gastrointestinal tract, e.g. insulin or leuprolide
XX
PS   Claim 49; Page 56; 294pp; English.
XX
CC   This sequence encodes a peptide that specifically binds to the human
CC   sucrose-isomaltase complex. The invention relates to purified
CC   proteins (I) that bind specifically to at least one of the
CC   gastro-intestinal (GI) tract receptors human intestinal
CC   peptide-associated transporter (HPT1), hPEPT1, D2H and human
CC   sucrose-isomaltase complex (hSI). (I) provide active transport of
CC   therapeutic agents through human and animal GI tissue (into the blood)
CC   for in vivo delivery, particularly for treatment or prevention of
CC   hypertension, diabetes, osteoporosis, haemophilia, anaemia, cancer,
CC   migraine, or angina pectoris. Specifically they are used to deliver
CC   insulin or leuprolide, but many other suitable therapeutic agents are
CC   disclosed, including genes or inhibitory nucleic acid, imaging agents and
CC   antigens. (I) may also provide targeting to the GI tract. Other uses of
CC   (I) are: (i) to determine the level of specified receptors in a sample
CC   (in a binding assay); and (ii) to screen for molecules that bind (I).
CC   Immunogenic analogues or derivatives of (I) are used to raise antibodies
CC   and in immunoassays. The antibodies are used to locate, detect and
CC   measure (I), e.g. for imaging, monitoring treatment, tissue analysis
CC   etc., also for peptide purification and immobilisation.
XX
SQ   Sequence 177 BP; 31 A; 51 C; 53 G; 42 T; 0 other;


  Query Match             0.6%;  Score 18;  DB 20;  Length 177;
  Best Local Similarity   100.0%;  Pred. No. 58;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1816 agcagtaccacaaccagt 1833
           ||||||||||||||||||
Db      90 AGCAGTACCACAACCAGT 73


RESULT   8
T67773/c
ID   T67773 standard; DNA; 531 BP.
XX
AC   T67773;
XX
DT   29-JUL-1997  (first entry)
XX
DE   H. pylori cytoplasmic protein ORF 24824087.aa.
XX
KW   Vaccine; prevention; treatment; infection; identification;
KW   binding compound; bacterium; life cycle; activator; bacteria;
KW   inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis;
KW   cytoplasmic; ds.
XX
OS   Helicobacter pylori.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..531
FT                   /*tag=  a
FT                   /transl_except= (pos: 460..462, aa: Xaa)
FT                   /transl_except= (pos: 520..522, aa: Xaa)
FT                   /transl_except= (pos: 526..528, aa: Xaa)
FT                   /note= "Xaa = Unknown"
XX
PN   WO9640893-A1.
XX
PD   19-DEC-1996.
XX
PF   06-JUN-1996;   96WO-US09122.
XX
PR   01-APR-1996;   96US-0630405.
PR   07-JUN-1995;   95US-0487032.
XX
PA   (ASTR ) ASTRA AB.
XX
PI   Berglindh OT,  Smith D,  Mellgaerd BL;
XX
DR   WPI; 1997-052306/05.
DR   P-PSDB; W20335.
XX
PT   Helicobacter pylori nucleic acid sequences and related
PT   polypeptide(s)  - useful for vaccines to treat or prevent H. pylori
PT   infection, and to detect Helicobacter
XX
PS   Claim 9; Page -; 148lpp; English.
XX
CC   The present sequence encodes a Helicobacter pylori cytoplasmic
CC   protein.
CC   The protein may be used in a vaccine to prevent or treat H. pylori
CC   infection or to identify H. pylori polypeptide binding compounds,
CC   useful as potential H. pylori life cycle activators or inhibitors.
CC   The genomic sequence of H. pylori (ATCC 55679) was determined from
CC   overlapping contigs generated by mechanically shearing the
CC   bacterial DNA. The sequences were analysed for ORF of at least 180
CC   nucleotides, and the predicted coding regions defined by computer
CC   evaluation. To identify likely H. pylori antigens for vaccine
CC   development, the amino acid sequences predicted from various ORF
CC   were analysed for significant homology to other known or exported
CC   membrane proteins. Having identified and determined the sequences
CC   of interest, particular regions can be isolated from H. pylori by
CC   PCR amplification for recombinant polypeptide production, e.g. in
CC   E. coli hosts.
CC   Note: This DNA sequence is not reproduced in the specification and
CC   has been derived from the related specification, WO9719098.
XX
SQ   Sequence 531 BP; 175 A; 88 C; 114 G; 151 T; 3 other;


  Query Match             0.6%;  Score 18;  DB 18;  Length 531;
  Best Local Similarity   100.0%;  Pred. No. 57;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1760 aaagccgcttcacccttt 1777
           ||||||||||||||||||
Db     311 AAAGCCGCTTCACCCTTT 294


RESULT   9
T77453/c
ID   T77453 standard; DNA; 531 BP.
XX
AC   T77453;
XX
DT   11-AUG-1997  (first entry)
XX
DE   H. pylori cytoplasmic protein ORF 24824087.aa.
XX
KW   Chronic gastritis; duodenal ulcer disease; activator;
KW   inhibitor; bacterial life cycle; vaccine; immunisation; detection;
KW   antisense; inhibition; cytoplasmic; Na+/H+ antiporter;
KW   Escherichia coli; ds.
XX
OS   Helicobacter pylori.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..531
FT                   /*tag=  a
FT                   /transl_except= (pos: 460..462, aa: Xaa)
FT                   /transl_except= (pos: 520..522, aa: Xaa)
FT                   /transl_except= (pos: 526..528, aa: Xaa)
FT                   /note= "Xaa = Unknown"
XX
PN   WO9719098-A1.
XX
```

PD 29-MAY-1997.
XX
PF 15-NOV-1996; 96WO-US18542.
XX
PR 17-NOV-1995; 95US-0561469.
XX
PA (ASTR) ASTRA AB.
XX
PI Smith DH;
XX
DR WPI; 1997-298052/27.
DR P-PSDB; W24635.
XX
PT Helicobacter pylori nucleic acid sequences and related proteins -
PT used for diagnostics and therapeutics
XX
PS Claim 1; Page 100; 235pp; English.
XX
CC The present sequence encodes a Helicobacter pylori cytoplasmic
CC protein, which was found to be homologous to Escherichia coli
CC Na+/H+ antiporter protein following BLAST protein analysis.
CC H. pylori has been strongly linked to chronic gastritis and
CC duodenal ulcer disease. The nucleic acid sequences of the invention
CC are used to evaluate compounds, especially activators or inhibitors
CC of bacterial life cycle, for the ability to bind an H. pylori
CC nucleic acid sequence. The nucleic acid sequences, and
CC corresponding proteins, are also useful for generating vaccines for
CC immunising subjects against H. pylori or for use in detecting the
CC presence of Helicobacter species in a sample. Antisense nucleic
CC acid sequences of these sequences are used to inhibit expression of
CC a gene from Helicobacter species. H. pylori whole genomic DNA was
CC isolated and nebulised to a median size of 2000 bp. Purified DNA
CC fragments were blunt-ended and ligated to unique BstXI-linker
CC adapters in 100-1000 fold molar excess. These linkers are
CC complementary to the BstXI-cut pMPX vectors, while the overhang is
CC not self-complementary. Therefore the linkers will not
CC concatemerise nor will the cut vector re-ligate itself easily. The
CC linker-adapter inserts were ligated to each of the 20 pMPX vectors
CC to construct a series of shotgun subclone libraries. The purified
CC DNA samples were then sequenced.
CC Note: The ORF/protein reference number for this sequence was
CC obtained from the related specification, WO9640893.
XX
SQ Sequence 531 BP; 175 A; 88 C; 114 G; 151 T; 3 other;

Query Match 0.6%; Score 18; DB 18; Length 531;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1760 aaagccgcttcacccttt 1777
||||||||||||||||||
Db 311 AAAGCCGCTTCACCCTTT 294

RESULT 10
Z80749
ID Z80749 standard; cDNA; 726 BP.
XX
AC Z80749;
XX
DT 07-APR-2000 (first entry)
XX
DE Human colon cancer cell line SW480 cDNA clone SEQ ID NO:833.
XX
KW Human; gene expression product; diagnosis; tumour; colon cancer;
KW colorectal adenocarcinoma; cell line SW480; cell proliferation;
KW cytostatic; sarcoma; breast cancer; neoplasia; dysplasia;
KW hyperplasia; ds.
XX
OS Homo sapiens.
XX
PN WO9964576-A2.
XX
PD 16-DEC-1999.
XX
PF 09-JUN-1999; 99WO-IB01062.
XX
PR 10-JUN-1998; 98US-0088801.
XX
PA (FARB) BAYER CORP.
XX
PI Endege WO, Steinmann KE, Astle JH, Burgess CC, Bushnell SE;
PI Carroll E, Catino TJ, Derti A, Ford DM, Lewis ME, Monahan JE;
PI Schlegel R;
XX
DR WPI; 2000-087220/07.
XX
PT Novel nucleic acids, used to develop products for the diagnosis and
PT treatment of disorders involving unwanted cell proliferation,
PT particularly cancers, especially colon cancer -
XX
PS Claim 15; Page 462; 469pp; English.
XX
CC Z79917 to Z80766 represent double stranded cDNA clones isolated from the
CC human colorectal adenocarcinoma (colon cancer) cell line SW480. The
CC cDNA clones can be used to generate antisense oligonucleotides which
CC can be used for antisense therapy. Methods and products from the present
CC invention can be used for identifying and/or classifying cancerous cells
CC present in a human tumour, particularly in solid tumours, e.g. carcinomas
CC and sarcomas, e.g. breast or colon cancers. The cDNA clones can be used
CC for developing agents for the diagnosis and treatment of disorders
CC involving unwanted cell proliferation, such as neoplasia, dysplasia or
CC hyperplasia.
XX
SQ Sequence 726 BP; 149 A; 157 C; 170 G; 198 T; 52 other;

Query Match 0.6%; Score 18; DB 21; Length 726;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 227 cctgcaggtggtggcagc 244
||||||||||||||||||
Db 60 cctgcaggtggtggcagc 77

RESULT 11
Z48812
ID Z48812 standard; cDNA; 1040 BP.
XX
AC Z48812;
XX
DT 21-MAR-2000 (first entry)
XX
DE Soybean inositol 1,3,4-triphosphate 5/6-kinase coding sequence.
XX
KW Inositol 1,3,4-triphosphate 5/6-kinase; phytic acid; genetic mapping;
KW myo-inositol 1,2,3,4,5,6-hexaphosphate; biosynthetic enzyme; phytate;
KW animal feed; ss.
XX
OS Glycine max.
XX
PN WO9955879-A1.
XX
PD 04-NOV-1999.
XX
PF 22-APR-1999; 99WO-US08790.
XX
PR 24-APR-1998; 98US-0082960.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Cahoon RE, Carlson TJ, Hitz WD, Pearlstein RW;
XX
DR WPI; 2000-072179/06.

```
DR   P-PSDB; Y59428.
XX
PT   Novel phytic acid biosynthetic enzymes used to alter the level of the
PT   enzyme in transgenic plants  -
XX
PS   Claim 2; Page 53; 63pp; English.
XX
CC   This sequence encodes an inositol 1,3,4-triphosphate 5/6-kinase (ITK)
CC   protein of the invention. ITK is a phytic acid (Myo-inositol
CC   1,2,3,4,5,6-hexaphosphate) biosynthetic enzyme. The ITK enzymes of the
CC   invention may be prepared recombinantly and used to raise antibodies,
CC   which are used for detecting the enzymes in situ in cells or in vitro
CC   in cell extracts. The polynucleotides of the invention may be used to
CC   create transgenic plants in which the ITK levels are present at higher
CC   or lower levels than normal, or in cell types or developmental processes
CC   where they are not normally found. This would alter the level of
CC   1,3,4-triphosphate 5/6 kinase found in those cells. In addition, it may
CC   be desirable to eliminate expression of ITK genes for certain
CC   applications. The polynucleotides also serve as a source of probes and
CC   primers, which are useful for genetic mapping, as markers for traits
CC   linked to those genes, and to isolate homologous sequences from other
CC   species, as well as for physical mapping. Decreased amounts of phytate
CC   and increased amounts of available phosphate in animal feeds such as corn
CC   would lead to improved feed efficiency. The proteins of the present
CC   invention lead to a better understanding of the phytic acid biosynthesis
CC   pathway, allowing it to be exploited for commercial uses, e.g. in animal
CC   feeds.
XX
SQ   Sequence 1040 BP; 264 A; 255 C; 265 G; 256 T; 0 other;


  Query Match             0.6%;  Score 18;  DB 21;  Length 1040;
  Best Local Similarity   100.0%;  Pred. No. 56;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    990 gaaagcttcattcaaccc 1007
          ||||||||||||||||||
Db     56 gaaagcttcattcaaccc 73


RESULT  12
T67992/c
ID   T67992 standard; DNA; 1593 BP.
XX
AC   T67992;
XX
DT   16-JUL-1997  (first entry)
XX
DE   H. pylori cytoplasmic protein ORF 06cp30603orf11.
XX
KW   Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW   identification; binding compound; bacterium; life cycle; activator;
KW   bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;
KW   diagnosis; ds.
XX
OS   Helicobacter pylori.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1593
FT                   /*tag=  a
FT                   /note= "no stop codon given"
XX
PN   WO9640893-A1.
XX
PD   19-DEC-1996.
XX
PF   06-JUN-1996;   96WO-US09122.
XX
PR   01-APR-1996;   96US-0630405.
PR   07-JUN-1995;   95US-0487032.
XX
PA   (ASTR ) ASTRA AB.
XX
PI   Berglindh OT,  Smith D,  Mellgaerd BL;
XX
DR   WPI; 1997-052306/05.
DR   P-PSDB; W20739.
XX
PT   Helicobacter pylori nucleic acid sequences and related
PT   polypeptide(s)  - useful for vaccines to treat or prevent H. pylori
PT   infection, and to detect Helicobacter
XX
PS   Claim 9; Page 833; 1481pp; English.
XX
CC   The present sequence encodes a H. pylori cytoplasmic protein.
CC   The protein may be used in a vaccine to prevent or treat H. pylori
CC   infection or to identify H. pylori polypeptide binding compounds,
CC   useful as potential H. pylori life cycle activators or inhibitors.
CC   The genomic sequence of H. pylori (ATCC 55679) was determined from
CC   overlapping contigs generated by mechanically shearing the bacterial
CC   DNA. The sequences were analysed for ORF of at least 180 nucleotides,
CC   and the predicted coding regions defined by computer evaluation. To
CC   identify likely H. pylori antigens for vaccine development, the amino
CC   acid sequences predicted from various ORF were analysed for significant
CC   homology to other known or exported membrane proteins. Having identified
CC   and determined the sequences of interest, particular regions can be
CC   isolated from H. pylori by PCR amplification for recombinant polypeptide
CC   production, e.g. in E. coli hosts.
XX
SQ   Sequence 1593 BP; 541 A; 288 C; 334 G; 430 T; 0 other;


  Query Match             0.6%;  Score 18;  DB 18;  Length 1593;
  Best Local Similarity   100.0%;  Pred. No. 56;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1760 aaagccgcttcacccttt 1777
          ||||||||||||||||||
Db   1376 AAAGCCGCTTCACCCTTT 1359


RESULT  13
X08683/c
ID   X08683 standard; cDNA; 1626 BP.
XX
AC   X08683;
XX
DT   27-SEP-1999  (first entry)
XX
DE   Novel nucleotide sequence encoding new protein (Clone AC222_1).
XX
KW   Polynucleotide; protein; nutrition; cytokine; cell proliferation;
KW   cell differentiation; immunostimulation; immunosuppression;
KW   haematopoiesis regulation; tissue growth; activin; inhibin;
KW   chemotaxis; chemokinesis; haemostasis; thrombolysis; receptor;
KW   ligand; anti-inflammatory; tumour suppression; gene therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             161..1129
FT                   /*tag=  a
FT                   /product= "Novel protein"
XX
PN   WO9920644-A1.
XX
PD   29-APR-1999.
XX
PF   16-OCT-1998;   98WO-US22034.
XX
PR   18-OCT-1997;   97US-0955557.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Agostino MJ,  Bowman MR,  Evans C,  Jacobs K,  Lavallie ER;
PI   McCoy JM,  Merberg D,  Racie LA,  Spaulding V,  Treacy M;
```

XX
RL WPI; 1999-288272/24.
DR P-PSDB; W85718.
XX
PT New polynucleotides encoding secreted human proteins
XX
PS Claim 14; Page 101-102; 136pp; English.
XX
CC The new human secreted proteins are encoded by polynucleotides
CC obtained from human placenta, adult testes, fetal kidney, fetal
CC brain, adult brain, adult brain and adult blood cDNA libraries.
CC The polynucleotides and proteins are predicted to have biological
CC activities which would make them suitable for treating, preventing or
CC ameliorating medical conditions in humans and animals. Suggested
CC activities include nutritional activity, cytokine and cell
CC proliferation/differentiation activity, immune stimulating (e.g. as
CC vaccines) or suppressing activity, haematopoiesis regulating
CC activity, tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
CC activity, receptor/ligand activity, anti-inflammatory activity,
CC cadherin/tumour invasion suppressor activity, and tumour inhibition
CC activity. The polynucleotides are also stated to be useful for gene
CC therapy. The sequences are identified by a secretory leader
CC sequence motif in the polynucleotide and it is thought that the
CC encoded proteins have biological activity by virtue of their secreted
CC nature. This clone was designated AC222_1. A probe for this clone is
CC described in X08698.
XX
SQ Sequence 1626 BP; 560 A; 327 C; 291 G; 443 T; 5 other;

Query Match 0.6%; Score 18; DB 20; Length 1626;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1342 gggagtggcagagggatg 1359
||||||||||||||||||
Db 930 GGGAGTGGCAGAGGGATG 913

RESULT 14
V17351
ID V17351 standard; DNA; 2186 BP.
XX
AC V17351;
XX
DT 04-JUN-1998 (first entry)
XX
DE Coding sequence for Kex2 protease inhibitor.
XX
KW Kex2 protease; inhibitor; pesticide; kexstatin; ds.
XX
OS Streptomyces platensis.
XX
PN EP814160-A2.
XX
PD 29-DEC-1997.
XX
PF 19-JUN-1997; 97EP-0110075.
XX
PR 03-MAR-1997; 97JP-0048101.
PR 19-JUN-1996; 96JP-0158677.
PR 26-AUG-1996; 96JP-0224104.
XX
PA (BIOM-) BIOMOLECULAR ENG RES INST.
XX
PI Kikuchi N, Oda K, Shibano Y;
XX
DR WPI; 1998-044337/05.
XX
PT Kex2 protease inhibitor protein - with potential pharmaceutical and
PT pesticidal applications
XX
PS Disclosure; Page 18-19; 29pp; English.
XX
CC This sequence is the coding sequence for an example of the inhibitor of
CC the invention (the encoded protein is not given in the specification).
CC The inhibitor is a Kex2 proteinase family enzyme inhibitor with a
CC molecular weight of 11,500. The inhibitor/protein (termed kexstatin) is
CC expected to have pharmaceutical and pesticidal applications.
XX
SQ Sequence 2186 BP; 297 A; 827 C; 780 G; 282 T; 0 other;

Query Match 0.6%; Score 18; DB 19; Length 2186;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1652 ggtcctgggcaccctggc 1669
||||||||||||||||||
Db 1518 ggtcctgggcaccctggc 1535

RESULT 15
X20537/c
ID X20537 standard; DNA; 2450 BP.
XX
AC X20537;
XX
DT 05-MAY-1999 (first entry)
XX
DE Polynucleotide sequence from the genome of Treponema pallidum.
XX
KW Treponema pallidum infection; syphilis; Borrelia infection; animal;
KW enzyme production; ds.
XX
OS Treponema pallidum.
XX
PN WO9859034-A2.
XX
PD 30-DEC-1998.
XX
PF 23-JUN-1998; 98WO-US13041.
XX
PR 24-JUN-1997; 97US-0050667.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Fraser CM;
XX
DR WPI; 1999-081273/07.
XX
PT New isolated Treponema pallidum nucleic acids - used to develop
PT products for the detection, diagnosis, characterisation, prevention
PT and therapy of T. pallidum infections, particularly syphilis
XX
PS Claim 1; Page 408-410; 1150pp; English.
XX
CC X20500-21243 represent polynucleotide sequences from the genome of
CC Treponema pallidum. The sequences can be used for detection,
CC diagnosis, characterisation, prevention and therapy for T. pallidum
CC infections, particularly syphilis. They can also be used for detecting
CC diseases related to Borrelia infections in animals, and for the
CC production of biosynthetic products such as enzymes.
XX
SQ Sequence 2450 BP; 553 A; 643 C; 631 G; 617 T; 6 other;

Query Match 0.6%; Score 18; DB 20; Length 2450;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2305 tctgctttggagactttc 2322
||||||||||||||||||
Db 941 TCTGCTTTGGAGACTTTC 924

ESULT 16
04699/c
D V04699 standard; cDNA; 2605 BP.
XX
AC V04699;
XX
DT 21-JUL-1998 (first entry)
XX
DE Homo sapiens 20q13 amplicon cc43 transcript.
XX
KW 20q13 amplicon; chromosome 20; tumour; detection;
KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
KW treatment; age-related macular degeneration; retinitis pigmentation;
KW Leber's congenital amaurosis; ds.
XX
OS Homo sapiens.
XX
PN WO9802539-A1.
XX
PD 22-JAN-1998.
XX
PF 15-JUL-1997; 97WO-US12343.
XX
PR 17-JAN-1997; 97US-0785532.
PR 15-JUL-1996; 96US-0680395.
PR 16-OCT-1996; 96US-0731499.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Collins CC, Godfrey T, Gray JW, Hwang SI, Kowbel D;
PI Rommens J;
XX
DR WPI; 1998-110587/10.
XX
PT New sequences from the 20q13 amplicon - used for detecting
PT chromosomal abnormalities, particularly tumours, and for developing
PT products for treating diseases
XX
PS Claim 1; Page 62; 91pp; English.
XX
CC The sequence is that of a cDNA sequence cc43, which was isolated
CC from the 20q13 amplicon. It is expressed in normal tissues but not
CC been found in the breast cancer cell line. It can be used as a probe
CC for the detection of chromosomal abnormalities at 20q13. It and
CC other sequences isolated from the 20q13 amplicon are consistently
CC amplified in primary tumours. These sequences are useful as probes
CC or as probe targets for monitoring the relative copy number of
CC corresponding sequences from a biological sample such as tumour
CC cells. The sequences can also be used in therapeutic applications
CC for modulating the expression of the endogenous gene or the activity
CC of the gene product. Examples of therapeutic approaches include
CC antisense inhibition of gene expression, gene therapy, and monoclonal
CC antibodies that specifically bind the gene products. The products can
CC also be used in the treatment of other diseases, e.g. age-related
CC macular degeneration, Leber's congenital amaurosis and retinitis
CC pigmentation.
XX
SQ Sequence 2605 BP; 698 A; 625 C; 549 G; 732 T; 1 other;

Query Match 0.6%; Score 18; DB 19; Length 2605;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 448 ttccaaagtgtgtacttt 465
|||||||||||||||||||
Db 1421 TTCCAAAGTGTGTACTTT 1404

ESULT 17
33895
D Z33895 standard; cDNA; 2945 BP.
XX
AC Z33895;
XX
DT 07-DEC-1999 (first entry)
XX
DE Human PRO274 nucleotides sequence.
XX
KW Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
KW probe; blood coagulation disorder; cancer; cellular adhesion disorder;
KW secreted protein; transmembrane protein; ss.
XX
OS Homo sapiens.
XX
PN WO9946281-A2.
XX
PD 16-SEP-1999.
XX
PF 08-MAR-1999; 99WO-US05028.
XX
PR 10-MAR-1998; 98US-0077450.
PR 11-MAR-1998; 98US-0077632.
PR 11-MAR-1998; 98US-0077641.
PR 11-MAR-1998; 98US-0077649.
PR 12-MAR-1998; 98US-0077791.
PR 13-MAR-1998; 98US-0078004.
PR 17-MAR-1998; 98US-0040220.
PR 20-MAR-1998; 98US-0078886.
PR 20-MAR-1998; 98US-0078910.
PR 20-MAR-1998; 98US-0078936.
PR 20-MAR-1998; 98US-0078939.
PR 25-MAR-1998; 98US-0079294.
PR 26-MAR-1998; 98US-0079656.
PR 27-MAR-1998; 98US-0079663.
PR 27-MAR-1998; 98US-0079664.
PR 27-MAR-1998; 98US-0079689.
PR 27-MAR-1998; 98US-0079728.
PR 27-MAR-1998; 98US-0079786.
PR 30-MAR-1998; 98US-0079920.
PR 30-MAR-1998; 98US-0079923.
PR 31-MAR-1998; 98US-0080105.
PR 31-MAR-1998; 98US-0080107.
PR 31-MAR-1998; 98US-0080165.
PR 31-MAR-1998; 98US-0080194.
PR 01-APR-1998; 98US-0080327.
PR 01-APR-1998; 98US-0080328.
PR 01-APR-1998; 98US-0080333.
PR 01-APR-1998; 98US-0080334.
PR 08-APR-1998; 98US-0081049.
PR 08-APR-1998; 98US-0081070.
PR 08-APR-1998; 98US-0081071.
PR 09-APR-1998; 98US-0081195.
PR 09-APR-1998; 98US-0081203.
PR 09-APR-1998; 98US-0081229.
PR 15-APR-1998; 98US-0081817.
PR 15-APR-1998; 98US-0081838.
PR 15-APR-1998; 98US-0081952.
PR 15-APR-1998; 98US-0081955.
PR 21-APR-1998; 98US-0082568.
PR 21-APR-1998; 98US-0082569.
PR 22-APR-1998; 98US-0082700.
PR 22-APR-1998; 98US-0082704.
PR 22-APR-1998; 98US-0082804.
PR 23-APR-1998; 98US-0082767.
PR 23-APR-1998; 98US-0082796.
PR 27-APR-1998; 98US-0083336.
PR 28-APR-1998; 98US-0083322.
PR 29-APR-1998; 98US-0083392.
PR 29-APR-1998; 98US-0083495.
PR 29-APR-1998; 98US-0083496.
PR 29-APR-1998; 98US-0083499.
PR 29-APR-1998; 98US-0083500.
PR 29-APR-1998; 98US-0083545.
PR 29-APR-1998; 98US-0083554.

```
:R   29-APR-1998;   98US-0083558.
:R   29-APR-1998;   98US-0083559.
:R   30-APR-1998;   98US-0083742.
:R   05-MAY-1998;   98US-0084366.
:R   06-MAY-1998;   98US-0084414.
:R   06-MAY-1998;   98US-0084441.
:R   07-MAY-1998;   98US-0084598.
:R   07-MAY-1998;   98US-0084600.
:R   07-MAY-1998;   98US-0084627.
:R   07-MAY-1998;   98US-0084637.
:R   07-MAY-1998;   98US-0084639.
:R   07-MAY-1998;   98US-0084640.
:R   07-MAY-1998;   98US-0084643.
:R   13-MAY-1998;   98US-0085323.
:R   13-MAY-1998;   98US-0085338.
:R   13-MAY-1998;   98US-0085339.
:R   15-MAY-1998;   98US-0085573.
:R   15-MAY-1998;   98US-0085579.
:R   15-MAY-1998;   98US-0085580.
:R   15-MAY-1998;   98US-0085582.
:R   15-MAY-1998;   98US-0085689.
:R   15-MAY-1998;   98US-0085697.
:R   15-MAY-1998;   98US-0085700.
:R   15-MAY-1998;   98US-0085704.
:R   18-MAY-1998;   98US-0086023.
:R   22-MAY-1998;   98US-0086392.
:R   22-MAY-1998;   98US-0086414.
:R   22-MAY-1998;   98US-0086430.
:R   22-MAY-1998;   98US-0086486.
:R   28-MAY-1998;   98US-0087098.
:R   28-MAY-1998;   98US-0087106.
:R   28-MAY-1998;   98US-0087208.
:R   30-JUL-1998;   98US-0094651.
:R   11-SEP-1998;   98US-0100038.
:X
:A   (GETH ) GENENTECH INC.
:X
:I   Wood WI,  Goddard A,  Gurney A,  Yuan J,  Baker KP,  Chen J;
:X
:R   WPI; 1999-551358/46.
:R   P-PSDB; Y41686.
:X
:T   New secreted and transmembrane polypeptides and their polynucleotides,
:T   useful for treating blood coagulation disorders, cancers and cellular
:T   adhesion disorders -
:X
:S   Claim 2; Fig 3; 530pp; English.
:X
:C   The present invention describes secreted and transmembrane polypeptides
:C   and their polynucleotides. The nucleotide sequences are useful as
:C   sources of probes, primers, for chromosome mapping, and for generation
:C   of antisense sequences. They can also be used to create transgenic
:C   animals. The proteins can be used to treat a variety of diseases and
:C   disorders, depending on their function. Diseases that may be treated
:C   include blood coagulation disorders, cancers and cellular adhesion
:C   disorders. They may also be used to raise antibodies. Z33891 to
:C   Z34338, and Y41685 to Y41774 represent polynucleotide and polypeptide
:C   sequence given in the exemplification of the present invention.
:X
:Q   Sequence 2945 BP; 703 A; 776 C; 705 G; 761 T; 0 other;

  Query Match             0.6%;  Score 18;  DB 20;  Length 2945;
  Best Local Similarity   100.0%;  Pred. No. 55;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    872 cattgctgctgtcaagga 889
          ||||||||||||||||||
Db    193 cattgctgctgtcaagga 210


RESULT  18
Z09473/c
ID   Z09473 standard; DNA; 4120 BP.
XX
AC   Z09473;
XX
DT   02-NOV-1999  (first entry)
XX
DE   Human RNA helicase p135 DNA #2.
XX
KW   DEAD protein; insect cell vector; DEAD-superfamily; RNA helicase;
KW   DNA helicase; cancer research; cell line; inflammation; apoptosis;
KW   drug; anticancer; antiviral; p135; ds.
XX
OS   Homo sapiens.
XX
PN   WO9941390-A2.
XX
PD   19-AUG-1999.
XX
PF   01-FEB-1999;   99WO-EP00829.
XX
PR   12-FEB-1998;   98DE-1005781.
XX
PA   (AVET ) AVENTIS RES & TECHNOLOGIES GMBH & CO KG.
XX
PI   Boehnisch B,  Gallert K,  Huels C,  Muellner S;
XX
DR   WPI; 1999-527373/44.
XX
PT   New insect cell vector containing the sequence encoding a
PT   DEAD-superfamily protein, particularly a nucleic acid helicase, used
PT   e.g. for identifying potential pharmaceuticals
XX
PS   Disclosure; Page 38-40; 43pp; German.
XX
CC   This invention describes the construction of a novel insect cell vector
CC   (A) which contains a nucleic acid (I) that codes for a protein (II) of
CC   the DEAD-superfamily. (A), and recombinant insect viruses derived from
CC   them, are used to express recombinant (II), particularly RNA and DNA
CC   helicases. (II) are potentially useful for: (a) production of cell lines
CC   for research into cancer, inflammation and apoptosis, or for clarifying
CC   the mechanism of action of drugs, and (b) to identify pharmaceutical
CC   activity in known compounds, e.g. anticancer and antiviral activities.
CC   (II), which are difficult to express in bacteria and yeast, are expressed
CC   at high level in insect cells, e.g. 300-400 mg per 10⁹ cells. This
CC   sequence encodes a human RNA-helicase p135 protein which is used in the
CC   description of the method of the invention.
XX
SQ   Sequence 4120 BP; 1099 A; 1021 C; 1236 G; 764 T; 0 other;

  Query Match             0.6%;  Score 18;  DB 20;  Length 4120;
  Best Local Similarity   100.0%;  Pred. No. 55;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1338 aggagggagtggcagagg 1355
          ||||||||||||||||||
Db    809 AGGAGGGAGTGGCAGAGG 792


RESULT  19
V70354
ID   V70354 standard; DNA; 6139 BP.
XX
AC   V70354;
XX
DT   09-FEB-1999  (first entry)
XX
DE   Coding strand of native genomic hK2.
XX
KW   Prostate cancer; detection; hK2; hK1; hK3; phK2; tissue kallikrein;
KW   pphK2; prostate-specific glandular kallikrein protein; PSA; human;
KW   prostate-specific antigen; ss.
XX
```

```
OS   Homo sapiens.
XX
PN   WO9846795-A1.
XX
PD   22-OCT-1998.
XX
PF   09-APR-1998;   98WO-US07027.
XX
PR   11-APR-1997;   97US-0843076.
XX
PA   (BAYU ) BAYLOR COLLEGE MEDICINE.
PA   (MAYO-) MAYO FOUNDATION.
XX
PI   Slawin KM,  Tindall DJ,  Young CYF;
XX
DR   WPI; 1998-594592/50.
XX
PT   Detection of human kallikrein 2 RNA - by reverse transcription and
PT   amplification by PCR, for detecting, monitoring and staging of
PT   prostate cancer
XX
PS   Disclosure; Page 78-80; 90pp; English.
XX
CC   The present invention describes a diagnostic method for detecting human
CC   kallikrein 2 (hK2) DNA. The method comprises: (a) contacting DNA obtained
CC   by reverse transcription (RT) of RNA from a human physiological sample
CC   which comprises cells suspected of containing hK2 RNA with at least 2
CC   oligonucleotides to amplify the DNA by PCR to yield amplified hK2 DNA,
CC   where the conditions amplify the DNA obtained by RT of RNA from at least
CC   one cell containing hK2 in a sample which comprises at least 107 to 109
CC   cells; and (b) detecting the presence of the amplified hK2 DNA. The
CC   method can be used for detecting, monitoring the progression of and
CC   pathologically staging prostate cancer. The present sequence represents
CC   the coding strand of native genomic hK2.
XX
SQ   Sequence 6139 BP; 1233 A; 1875 C; 1595 G; 1436 T; 0 other;


  Query Match             0.6%;  Score 18;  DB 19;  Length 6139;
  Best Local Similarity   100.0%;  Pred. No. 54;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1048 cagatgcccccgtggcct 1065
            ||||||||||||||||||
Db   3041 cagatgcccccgtggcct 3058


RESULT  20
Z23903/c
ID   Z23903 standard; DNA; 49999 BP.
XX
AC   Z23903;
XX
DT   25-JAN-2000  (first entry)
XX
DE   Human LOBO homologue genomic DNA fragment 5.
XX
KW   LOBO; long bones; bone development; bone extension; skull; osteopathic;
KW   diagnostic; pharmaceutical; gene therapy; transgenic animal; disease;
KW   spondyloepiphysal dysplasia; achondroplasia; human; ds.
XX
OS   Homo sapiens.
XX
PN   WO9950284-A2.
XX
PD   07-OCT-1999.
XX
PF   26-MAR-1999;   99WO-EP02055.
XX
PR   27-MAR-1998;   98DE-1013799.
XX
PA   (ROSE/) ROSENTHAL A.
PI   Rosenthal A,  Rump A,  Hess J,  Aigner T,  Wirth T;
XX
DR   WPI; 1999-601320/51.
XX
PT   Nucleic acids encoding proteins which influence bone development,
PT   useful for treating and studying bone disorders -
XX
PS   Example 3; Page 328-356; 391pp; German.
XX
CC   This invention describes novel nucleic acids (I; designated LOBO (long
CC   bones)) encoding proteins influencing bone development in mammals. The
CC   proteins of the invention reduce and/or inactivate bone extension (i.e.
CC   development), with exception of the skull and have osteopathic activity.
CC   The nucleic acid molecules, proteins and antibodies can be used in
CC   diagnostic or pharmaceutical compounds e.g. for gene therapy. The methods
CC   and nucleic acid molecules, etc. are useful for production of transgenic
CC   animals, especially a transgenic mouse for the study of diseases
CC   associated with bone development, e.g. spondyloepiphysal dysplasia and
CC   achondroplasia. This sequence encodes a human LOBO protein described
CC   in the method of the invention.
XX
SQ   Sequence 49999 BP; 10983 A; 13723 C; 13439 G; 11854 T; 0 other;


  Query Match             0.6%;  Score 18;  DB 20;  Length 49999;
  Best Local Similarity   100.0%;  Pred. No. 52;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2740 gggccaggaggctgccac 2757
            ||||||||||||||||||
Db  14820 GGGCCAGGAGGCTGCCAC 14803


RESULT  21
Q54684
ID   Q54684 standard; DNA; 21 BP.
XX
AC   Q54684;
XX
DT   05-JUL-1994  (first entry)
XX
DE   C kappa exon primer.
XX
KW   Germ line; stem cell; blastocyst; implantation; embryonic;
KW   electroporation; kappa; constant; exon; ss.
XX
OS   Synthetic.
XX
PN   DE4228162-C.
XX
PD   13-JAN-1994.
XX
PF   25-AUG-1992;   92DE-4228162.
XX
PR   25-AUG-1992;   92DE-4228162.
XX
PA   (RAJE/) RAJEWSKY K.
PA   (KOEL-) KOELNER VER FOERDERUNG IMMUNOLOGIE.
XX
PI   Rajewsky K,  Zou Y;
XX
DR   WPI; 1994-008862/02.
XX
PT   Homologous replacement of gene in mammalian germ line - by
PT   transfecting embryonic stem cell with labelled recombination
PT   vehicle, selection and incorporating into blastocyst(s) for
PT   implantation, esp. for prodn. of humanised antibodies in mice
XX
PS   Example; Column 4; 7pp; German.
XX
CC   DSM 7211 is made by (1) inserting a fragment of pC-2 (contg.
CC   the C kappa exon) into pTZ-19(R); (2) inserting a 1.2 kb
CC   fragment of pHBc kappa (contg. the intron enhancer element);
```

CC (3) inserting a 1.1 kb fragment contg. the neomycin resistance
CC gene of pMCIneo; (4) inserting the 2.8 kb fragment of pHJkappa
CC (contg. the J1-5) elements); then (5) finally inserting a 427 bp
CC fragment made by PCR on mouse germ line DNA using one primer
CC (Q54683) which hybridises 12 bp after the start of the C kappa
CC exon and another (Q54684) which hybridises 137 bp after the stop
CC codon of this exon.
CC DSM 7211 was introduced into embryonic mouse stem cells by
CC electroporation. Wild type mice have in their serum no antibodies
CC carrying the human kappa gene constant region, but in mice
CC homozygotic for the mutation, such antibodies were present at
CC 2860 microg/ml.
XX
SQ Sequence 21 BP; 8 A; 2 C; 10 G; 1 T; 0 other;

Query Match 0.6%; Score 17; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1031 gaggtaccaaggaaagg 1047
|||||||||||||||||
Db 1 gaggtaccaaggaaagg 17

RESULT 22
A32079/c
ID A32079 standard; DNA; 256 BP.
XX
AC A32079;
XX
DT 05-JUL-2000 (first entry)
XX
DE Plant microsatellite marker #1040.
XX
KW Plant microsatellite sequence; core repeat sequence; detection; probe;
KW DNA polymorphism; genome mapping; physical mapping; fingerprinting;
KW variety identification; genetic variability evaluation; primer; ss.
XX
OS Pinus radiata.
XX
PN WO9967421-A1.
XX
PD 29-DEC-1999.
XX
PF 25-JUN-1999; 99WO-NZ00092.
XX
PR 25-JUN-1998; 98US-0105307.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Havukkala IJ, Bloksberg LN, Glenn M;
XX
DR WPI; 2000-116958/10.
XX
PT New plant microsatellite markers and associated flanking species for
PT the detection of polymorphic genetic markers -
XX
PS Claim 1; Page 382; 392pp; English.
XX
CC Sequences A31040-A32093 represent novel plant microsatellite sequences
CC and associated flanking species. The sequences comprise a central core
CC repeat sequence, especially selected from the sequences A32094-A32096
CC with left and right flanking sequences. The polynucleotide sequences
CC can be used in the detection of DNA polymorphisms, in genome mapping,
CC in physical mapping, in positional cloning of genes, in variety
CC identification and in evaluation of genetic variability within and
CC between plant tissues, populations, cultivars, species and species
CC groups. They may also be used to design hybridization probes for
CC oligonucleotide fingerprinting and library screening and to design
CC primers for microsatellite-primed PCR. Microsatellite markers are
CC useful to locate specific economically useful genes in plant genomes.
XX
SQ Sequence 256 BP; 89 A; 41 C; 35 G; 91 T; 0 other;

Query Match 0.6%; Score 17; DB 21; Length 256;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2021 tggctggaaagtggtct 2037
|||||||||||||||||
Db 256 TGGCTGGAAAGTGGTCT 240

RESULT 23
A32050/c
ID A32050 standard; DNA; 300 BP.
XX
AC A32050;
XX
DT 05-JUL-2000 (first entry)
XX
DE Plant microsatellite marker #1011.
XX
KW Plant microsatellite sequence; core repeat sequence; detection; probe;
KW DNA polymorphism; genome mapping; physical mapping; fingerprinting;
KW variety identification; genetic variability evaluation; primer; ss.
XX
OS Pinus radiata.
XX
PN WO9967421-A1.
XX
PD 29-DEC-1999.
XX
PF 25-JUN-1999; 99WO-NZ00092.
XX
PR 25-JUN-1998; 98US-0105307.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Havukkala IJ, Bloksberg LN, Glenn M;
XX
DR WPI; 2000-116958/10.
XX
PT New plant microsatellite markers and associated flanking species for
PT the detection of polymorphic genetic markers -
XX
PS Claim 1; Page 373; 392pp; English.
XX
CC Sequences A31040-A32093 represent novel plant microsatellite sequences
CC and associated flanking species. The sequences comprise a central core
CC repeat sequence, especially selected from the sequences A32094-A32096
CC with left and right flanking sequences. The polynucleotide sequences
CC can be used in the detection of DNA polymorphisms, in genome mapping,
CC in physical mapping, in positional cloning of genes, in variety
CC identification and in evaluation of genetic variability within and
CC between plant tissues, populations, cultivars, species and species
CC groups. They may also be used to design hybridization probes for
CC oligonucleotide fingerprinting and library screening and to design
CC primers for microsatellite-primed PCR. Microsatellite markers are
CC useful to locate specific economically useful genes in plant genomes.
XX
SQ Sequence 300 BP; 99 A; 45 C; 43 G; 113 T; 0 other;

Query Match 0.6%; Score 17; DB 21; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2021 tggctggaaagtggtct 2037
|||||||||||||||||
Db 292 TGGCTGGAAAGTGGTCT 276

ESULT 24
A01382
ID A01382 standard; cDNA; 300 BP.
XX
AC A01382;
XX
DT 19-MAY-2000 (first entry)
XX
DE Human colon cancer cell line polynucleotide sequence SEQ ID NO:1373.
XX
KW Human; colon cancer; tumour; diagnosis; gene expression product;
KW probe; detection; cancerous state; metastasis; identification;
KW breast cancer; oestrogen receptor-positive breast cancer; therapy;
KW oestrogen receptor-negative breast cancer; lung cancer; ss.
XX
OS Homo sapiens.
XX
PN WO9958675-A2.
XX
PD 18-NOV-1999.
XX
PF 13-MAY-1999; 99WO-US10602.
XX
PR 14-MAY-1998; 98US-0085426.
PR 15-MAY-1998; 98US-0085537.
PR 15-MAY-1998; 98US-0085696.
PR 21-OCT-1998; 98US-0105234.
PR 27-OCT-1998; 98US-0105877.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
PI Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;
XX
DR WPI; 2000-126369/11.
XX
PT Polynucleotide library used to determine cancerous states of mammalian
PT cells -
XX
PS Claim 1; Page 511; 1097pp; English.
XX
CC A00010 to A02716 represent polynucleotides isolated from cDNA libraries
CC constructed from human colon cancer cell lines. The present invention
CC also describes a method of detecting differentially expressed genes
CC correlated with a cancerous state of a mammalian cell, comprising
CC detecting at least one differentially expressed gene product in a test
CC sample derived from a cell suspected of being cancerous, where detection
CC of the differentially expressed gene product is correlated with a
CC cancerous state of the cell from which the test sample was derived.
CC The polynucleotides sequences can be used in a method for detecting
CC differentially expressed genes correlated with a cancerous state of a
CC mammalian cell. The polynucleotides can also be used as probes for
CC detecting and mapping related genes. They can be used in diagnosis and
CC prognosis of diseases and disorders (e.g. identification of
CC pre-metastatic or metastatic cancerous states, stages of cancer, or
CC responsiveness of cancer to therapy). This is particularly for breast
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
CC negative breast cancer, lung cancer, and colon cancer.
XX
SQ Sequence 300 BP; 42 A; 104 C; 84 G; 70 T; 0 other;

Query Match 0.6%; Score 17; DB 21; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2633 ccaggaggcagctcagg 2649
||||||||||||||||
Db 188 ccaggaggcagctcagg 204

RESULT 25
T07238/c
ID T07238 standard; DNA; 304 BP.
XX
AC T07238;
XX
DT 05-JUL-1996 (first entry)
XX
DE Immunogen DNA from n-(ABCDE) hepatitis virus.
XX
KW Non-A, non-B, non-C, non-D, non-E hepatitis virus; n-(ABCDE);
KW immunogen; antibody; vaccine; phage library; ds.
XX
OS Non-A, non-B, non-C, non-D, non-E hepatitis virus MY 190 clone D30.
XX
PN WO9532290-A2.
XX
PD 30-NOV-1995.
XX
PF 17-MAY-1995; 95WO-US05980.
XX
PR 20-MAY-1994; 94US-0246986.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Kim JP, Purcell RH;
XX
DR WPI; 1996-020585/02.
XX
PT New non-A, -B, -C, -D and -E (n-(ABCDE)) hepatitis DNA libraries -
PT used to develop prods. for the detection, diagnosis, prevention and
PT treatment of n-(ABCDE) hepatitis.
XX
PS Disclosure; Page 101; 165pp; English.
XX
CC The sequence represents clone D30 which encodes an immunogenic
CC polypeptide associated with non-A, non-B, non-C, non-D, non-E
CC hepatitis virus (n-(ABCDE) virus) infection, and is obtained by
CC preparing a phage from a phage lambda gt11 library of MY 190 DNA
CC (ATCC 75284) n-(ABCDE) hepatitis serum, plating to form plaques,
CC and screening the phage plaques for the production of polypeptides
CC immunoreactive with n-(ABCDE) serum. n-(ABCDE) hepatitis virus
CC polypeptides can be used for the production or detection of
CC antibodies, and in vaccines. The antibodies can be used for
CC detection, diagnosis and in passive immunotherapy. The DNA can
CC also be used in detection and diagnosis, and as hybridisation
CC probes for identification of further n-(ABCDE) hepatitis coding
CC sequences. Culture systems producing the n-(ABCDE) polypeptides
CC can be used in screening studies.
XX
SQ Sequence 304 BP; 65 A; 93 C; 94 G; 52 T; 0 other;

Query Match 0.6%; Score 17; DB 17; Length 304;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 132 cgccgcgagcggccgcg 148
|||||||||||||||||
Db 22 CGCCGCGAGCGGCCGCG 6

RESULT 26
Z42643/c
ID Z42643 standard; cDNA; 330 BP.
XX
AC Z42643;
XX
DT 01-FEB-2000 (first entry)
XX

DE Human 5' EST isolated from a cDNA library SEQ ID NO:402.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
KW gene therapy; chromosome mapping; upstream regulatory sequence;
KW forensic; location; development; protein synthesis; stability;
KW regulation; identification; ss.
XX
OS Homo sapiens.
XX
PN WO9953051-A2.
XX
PD 21-OCT-1999.
XX
PF 09-APR-1999; 99WO-IB00712.
XX
PR 09-APR-1998; 98US-0057719.
PR 28-APR-1998; 98US-0069047.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-038446/03.
DR P-PSDB; Y65029.
XX
PT Novel secreted protein 5' expressed sequence tag sequences used in
PT diagnostic, forensic, gene therapy, and chromosome mapping procedures
XX
PS Claim 1; Page 382; 837pp; English.
XX
CC Z42265 to Z43075 represent novel 5' expressed sequence tag (EST)
CC sequences, corresponding to human secreted proteins. Y64651 to Y65438
CC represent the EST-related proteins corresponding to Z42265 to Z43052.
CC The 5' ESTs can be used for producing secreted human gene products.
CC They can be used to identify and isolate 5' untranslated regions (UTRs)
CC and upstream regulatory regions which control the location, development
CC stage, rate, and quantity of protein synthesis, as well as stability of
CC mRNA. The ESTs are also useful as probes for chromosome mapping, and to
CC obtain full length cDNA clones. The ESTs can also be used in forensic
CC procedures to identify individuals, or in diagnostic procedures to
CC identify individuals having genetic diseases resulting from abnormal
CC gene expression. The products may also be used in gene therapy protocols.
CC The nucleic acids encoding signal peptides can be used for directing
CC extracellular secretion of a polypeptide or the insertion of a
CC polypeptide into a membrane, or importing a polypeptide into a cell.
CC The proteins encoded by the EST sequences may be useful in treating a
CC variety of human conditions. Secreted proteins have therapeutic value,
CC and the identification of new secreted proteins is valuable. Z42249 to
CC Z42264 and Y64644 to Y64650 represent sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 330 BP; 76 A; 102 C; 86 G; 65 T; 1 other;

Query Match 0.6%; Score 17; DB 21; Length 330;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1219 acctcatccacccggac 1235
|||||||||||||||||
Db 48 ACCTCATCCACCCGGAC 32

RESULT 27
X24993/c
ID X24993 standard; cDNA; 332 BP.
XX
AC X24993;
XX
DT 05-JUL-1999 (first entry)
XX
DE Murine Bcl-2 interacting mediator of cell death Bim-S cDNA.
XX
KW Bim-S; Bcl-2 interacting mediator of cell death; apoptosis;
KW cell cycle; mouse; cancer; autoimmune disease;
KW degenerative disease; therapy; contraceptive; splice variant;
KW isoform; ss.
XX
OS Mus musculus.
XX
PN WO9914321-A1.
XX
PD 25-MAR-1999.
XX
PF 17-SEP-1998; 98WO-AU00772.
XX
PR 24-SEP-1997; 97AU-0009373.
PR 17-SEP-1997; 97AU-0009263.
XX
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
PI Puthalakath H, Strasser A;
XX
DR WPI; 1999-244030/20.
DR P-PSDB; W98154.
XX
PT New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
PT treatment
XX
PS Claim 3; Page 92; 145pp; English.
XX
CC The present sequence encodes the short form (S) of murine Bim, or
CC Bcl-2 interacting mediator of cell death (see W98154), a novel
CC member of the Bcl-2 family that is capable of inducing cell death
CC (apoptosis) and which acts as a 'death-ligand' for certain members
CC of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the
CC only Bcl-2 homology region which it encompasses is BH3. It is the
CC only BH3-only protein for which splice variants exist. These
CC result in the expression of a variety of isoforms, i.e. Bim-S,
CC Bim-L and Bim-EL (see W98154-56). cDNAs encoding these murine Bim
CC isoforms were obtained from a T lymphoma cDNA library using human
CC recombinant Bcl-2 protein. The murine Bim gene has been mapped to
CC chromosome 2 at bands F3-G. Human Bim-L and Bim-EL isoforms have
CC also been identified (see W98157-58). Binding the dynein light
CC chain was shown to regulate the pro-apoptotic activity of Bim.
CC Bim-S, the splice variant which does not bind to dynein light
CC chain, is a much more potent killer than either Bim-L or Bim-EL.
CC The invention provides variants (see W98159-68) of murine and human
CC Bim-L or Bim-EL that cannot bind, couple or otherwise associate
CC with a dynein light chain. The identification of Bim permits the
CC identification and rational design of a range of products for use
CC in therapy, diagnosis, antibody generation and involving modulation
CC of physiological cell death. These therapeutic molecules may act
CC as either antagonists or agonists of Bim's function and will be
CC useful in cancer, autoimmune or degenerative disease therapy.
CC Increased Bim expression or Bim activity is useful, e.g. for
CC treatment or prophylaxis in conditions such as cancer and deletion
CC of autoreactive lymphocytes in autoimmune disease. Decreased Bim
CC expression of Bim activity is useful in regulating inhibition or
CC prevention of cell death or degeneration such as under cytotoxic
CC conditions during e.g. gamma-irradiation and chemotherapy or during
CC HIV/AIDS or other viral infections, ischemia, myocardial infarction,
CC hypoxia, degenerative diseases or for prolonging the survival of
CC cells being transplanted for treatment of disease. Since Bim is
CC expressed in germ cells, modulating Bim expression or Bim activity
CC is useful, e.g. as a contraceptive or method of sterilization by
CC preventing generation of fertile sperm.
XX
SQ Sequence 332 BP; 87 A; 85 C; 91 G; 69 T; 0 other;

Query Match 0.6%; Score 17; DB 20; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    654 aggcctctcagcaggct 670
          |||||||||||||||||
Db     75 AGGCCTCTCAGCAGGCT 59


RESULT  28
T24629
ID   T24629 standard; cDNA to mRNA; 347 BP.
XX
AC   T24629;
XX
DT   07-OCT-1996  (first entry)
XX
DE   Human gene signature HUMGS06689.
XX
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
XX
OS   Homo sapiens.
XX
PN   WO9514772-A1.
XX
PD   01-JUN-1995.
XX
PF   11-NOV-1994;   94WO-JP01916.
XX
PR   12-NOV-1993;   93JP-0355504.
XX
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
XX
PI   Matsubara K,  Okubo K;
XX
DR   WPI; 1995-206931/27.
XX
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
XX
PS   Claim 1; Page 1655; 2245pp; Japanese.
XX
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in T19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
XX
SQ   Sequence 347 BP; 61 A; 102 C; 108 G; 75 T; 1 other;

  Query Match             0.6%;  Score 17;  DB 16;  Length 347;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1336 ccaggagggagtggcag 1352
           |||||||||||||||||
Db     15 ccaggagggagtggcag 31


RESULT  29
X41005/c
ID   X41005 standard; cDNA; 369 BP.
XX
AC   X41005;
XX
DT   18-JUN-1999  (first entry)
XX
DE   Human secreted protein 5' EST SEQ ID NO: 217.
XX
KW   Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW   forensic; gene therapy; chromosome mapping; signal peptide;
KW   upstream regulatory sequence; cytokine activity; cell proliferation;
KW   differentiation; haematopoiesis regulation; tissue growth regulation;
KW   reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW   thrombolytic; anti-inflammatory; tumour inhibition; ds.
XX
OS   Homo sapiens.
XX
PN   WO9906554-A2.
XX
PD   11-FEB-1999.
XX
PF   31-JUL-1998;   98WO-IB01238.
XX
PR   01-AUG-1997;   97US-0905134.
XX
PA   (GEST ) GENSET.
XX
PI   Duclert A,  Dumas Milne Edwards J,  Lacroix B;
XX
DR   WPI; 1999-153784/13.
DR   P-PSDB; Y12172.
XX
PT   New nucleic acids encoding human secreted proteins - obtained from
PT   cDNA libraries prepared from kidney, fetal kidney, dystrophic
PT   muscle, muscle and heart tissue
XX
PS   Claim 1; Page 314-315; 622pp; English.
XX
CC   X40826 to X41093 represent 5' expressed sequence tags (ESTs) for human
CC   secreted proteins, and encode the proteins given in Y01602 and
CC   Y11994 to Y12260, respectively. The proteins given represent the signal
CC   peptide and an N-terminal fragment of a secreted protein. The nucleic
CC   acid sequences can be used for producing secreted human gene products.
CC   They can also be used to develop products for diagnosis and therapy.
CC   The proteins obtained may have cytokine activity, cell
CC   proliferation/differentiation activity, haematopoiesis regulating
CC   activity, tissue growth regulating activity, reproductive hormone
CC   regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC   thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC   activity, tumour inhibition activity or other activities. The products
CC   can be used in forensic, gene therapy and chromosome mapping procedures.
CC   The sequences can also be used for obtaining corresponding promoter
CC   sequences. The nucleic acids encoding the signal peptide can be used
CC   for directing extracellular secretion of a polypeptide or the insertion
CC   of a polypeptide into a membrane, or importing a polypeptide into
CC   a cell.
XX
SQ   Sequence 369 BP; 57 A; 128 C; 98 G; 84 T; 2 other;

  Query Match             0.6%;  Score 17;  DB 20;  Length 369;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2501 gccacaggccaagaagg 2517
           |||||.|||||||||||
Db    320 GCCACAGGCCAAGAAGG 304


RESULT  30
X24994/c
ID   X24994 standard; cDNA; 422 BP.
XX
```

AC X24994;
XX
DT 05-JUL-1999 (first entry)
XX
DE Murine Bcl-2 interacting mediator of cell death Bim-L cDNA.
XX
KW Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
KW cell cycle; mouse; cancer; autoimmune disease;
KW degenerative disease; therapy; contraceptive; splice variant;
KW isoform; ss.
XX
OS Mus musculus.
XX
PN WO9914321-A1.
XX
PD 25-MAR-1999.
XX
PF 17-SEP-1998; 98WO-AU00772.
XX
PR 24-SEP-1997; 97AU-0009373.
PR 17-SEP-1997; 97AU-0009263.
XX
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
PI Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
PI Puthalakath H, Strasser A;
XX
DR WPI; 1999-244030/20.
DR P-PSDB; W98155.
XX
PT New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
PT treatment
XX
PS Claim 3; Page 94-95; 145pp; English.
XX
CC The present sequence encodes the long form (L) of murine Bim, or
CC Bcl-2 interacting mediator of cell death (see W98155), a novel
CC member of the Bcl-2 family that is capable of inducing cell death
CC (apoptosis) and which acts as a 'death-ligand' for certain members
CC of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the
CC only Bcl-2 homology region which it encompasses is BH3. It is the
CC only BH3-only protein for which splice variants exist. These
CC result in the expression of a variety of isoforms, i.e. Bim-S,
CC Bim-L and Bim-EL (see W98154-56). cDNAs encoding these murine Bim
CC isoforms were obtained from a T lymphoma cDNA library using human
CC recombinant Bcl-2 protein. The murine Bim gene has been mapped to
CC chromosome 2 at bands F3-G. Human Bim-L and Bim-EL isoforms have
CC also been identified (see W98157-58). Binding the dynein light
CC chain was shown to regulate the pro-apoptotic activity of Bim.
CC Bim-S, the splice variant which does not bind to dynein light
CC chain, is a much more potent killer than either Bim-L or Bim-EL.
CC The invention provides variants (see W98159-68) of murine and human
CC Bim-L or Bim-EL that cannot bind, couple or otherwise associate
CC with a dynein light chain. The identification of Bim permits the
CC identification and rational design of a range of products for use
CC in therapy, diagnosis, antibody generation and involving modulation
CC of physiological cell death. These therapeutic molecules may act
CC as either antagonists or agonists of Bim's function and will be
CC useful in cancer, autoimmune or degenerative disease therapy.
CC Increased Bim expression or Bim activity is useful, e.g. for
CC treatment or prophylaxis in conditions such as cancer and deletion
CC of autoreactive lymphocytes in autoimmune disease. Decreased Bim
CC expression of Bim activity is useful in regulating inhibition or
CC prevention of cell death or degeneration such as under cytotoxic
CC conditions during e.g. gamma-irradiation and chemotherapy or during
CC HIV/AIDS or other viral infections, ischemia, myocardial infarction,
CC hypoxia, degenerative diseases or for prolonging the survival of
CC cells being transplanted for treatment of disease. Since Bim is
CC expressed in germ cells, modulating Bim expression or Bim activity
CC is useful, e.g. as a contraceptive or method of sterilization by
CC preventing generation of fertile sperm.
XX
SQ Sequence 422 BP; 112 A; 116 C; 109 G; 85 T; 0 other;

Query Match 0.6%; Score 17; DB 20; Length 422;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 654 aggcctctcagcaggct 670
||||||||||||||||||
Db 75 AGGCCTCTCAGCAGGCT 59

RESULT 31
A32069/c
ID A32069 standard; DNA; 426 BP.
XX
AC A32069;
XX
DT 05-JUL-2000 (first entry)
XX
DE Plant microsatellite marker #1030.
XX
KW Plant microsatellite sequence; core repeat sequence; detection; probe;
KW DNA polymorphism; genome mapping; physical mapping; fingerprinting;
KW variety identification; genetic variability evaluation; primer; ss.
XX
OS Pinus radiata.
XX
PN WO9967421-A1.
XX
PD 29-DEC-1999.
XX
PF 25-JUN-1999; 99WO-NZ00092.
XX
PR 25-JUN-1998; 98US-0105307.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Havukkala IJ, Bloksberg LN, Glenn M;
XX
DR WPI; 2000-116958/10.
XX
PT New plant microsatellite markers and associated flanking species for
PT the detection of polymorphic genetic markers -
XX
PS Claim 1; Page 379; 392pp; English.
XX
CC Sequences A31040-A32093 represent novel plant microsatellite sequences
CC and associated flanking species. The sequences comprise a central core
CC repeat sequence, especially selected from the sequences A32094-A32095
CC with left and right flanking sequences. The polynucleotide sequences
CC can be used in the detection of DNA polymorphisms, in genome mapping,
CC in physical mapping, in positional cloning of genes, in variety
CC identification and in evaluation of genetic variability within and
CC between plant tissues, populations, cultivars, species and species
CC groups. They may also be used to design hybridization probes for
CC oligonucleotide fingerprinting and library screening and to design
CC primers for microsatellite-primed PCR. Microsatellite markers are
CC useful to locate specific economically useful genes in plant genomes.
XX
SQ Sequence 426 BP; 124 A; 81 C; 71 G; 150 T; 0 other;

Query Match 0.6%; Score 17; DB 21; Length 426;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2021 tggctggaaagtggtct 2037
||||||||||||||||||
Db 257 TGGCTGGAAAGTGGTCT 241

RESULT 32

32052/c
ID A32052 standard; DNA; 431 BP.
XX
AC A32052;
XX
DT 05-JUL-2000 (first entry)
XX
DE Plant microsatellite marker #1013.
XX
KW Plant microsatellite sequence; core repeat sequence; detection; probe;
KW DNA polymorphism; genome mapping; physical mapping; fingerprinting;
KW variety identification; genetic variability evaluation; primer; ss.
XX
OS Pinus radiata.
XX
PN WO9967421-A1.
XX
PD 29-DEC-1999.
XX
PF 25-JUN-1999; 99WO-NZ00092.
XX
PR 25-JUN-1998; 98US-0105307.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Havukkala IJ, Bloksberg LN, Glenn M;
XX
DR WPI; 2000-116958/10.
XX
PT New plant microsatellite markers and associated flanking species for
PT the detection of polymorphic genetic markers -
XX
PS Claim 1; Page 374; 392pp; English.
XX
CC Sequences A31040-A32093 represent novel plant microsatellite sequences
CC and associated flanking species. The sequences comprise a central core
CC repeat sequence, especially selected from the sequences A32094-A32096
CC with left and right flanking sequences. The polynucleotide sequences
CC can be used in the detection of DNA polymorphisms, in genome mapping,
CC in physical mapping, in positional cloning of genes, in variety
CC identification and in evaluation of genetic variability within and
CC between plant tissues, populations, cultivars, species and species
CC groups. They may also be used to design hybridization probes for
CC oligonucleotide fingerprinting and library screening and to design
CC primers for microsatellite-primed PCR. Microsatellite markers are
CC useful to locate specific economically useful genes in plant genomes.
XX
SQ Sequence 431 BP; 129 A; 78 C; 71 G; 152 T; 1 other;

Query Match 0.6%; Score 17; DB 21; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2021 tggctggaaagtggtct 2037
||||||||||||||||
Db 262 TGGCTGGAAAGTGGTCT 246

RESULT 33
V29360/c
ID V29360 standard; DNA; 559 BP.
XX
AC V29360;
XX
DT 31-JUL-1998 (first entry)
XX
DE Calcium ion channel alpha1 subunit exons 33, 34/intron; partial sequence.
XX
KW Calcium ion channel alpha1 subunit; human; episodic ataxia type 2;
KW familial hemiplegic migraine; FHM; EA-2; treatment; diagnosis;
KW exon; intron; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT intron 1..156
FT /*tag= a
FT /number= 32
FT /note= "partial sequence"
FT exon 157..222
FT /*tag= b
FT /number= 33
FT intron 223..394
FT /*tag= c
FT /number= 33
FT exon 395..509
FT /*tag= d
FT /number= 34
FT intron 510..559
FT /*tag= e
FT /number= 34
FT /note= "partial sequence"
XX
PN EP834561-A1.
XX
PD 08-APR-1998.
XX
PF 27-SEP-1996; 96EP-0202707.
XX
PR 27-SEP-1996; 96EP-0202707.
XX
PA (UYLE-) RIJKSUNIV LEIDEN.
XX
PI Ferrari MD, Frants RRIE, Ophoff RA, Terwindt GM;
XX
DR WPI; 1998-195461/18.
XX
PT New human nucleic acid associated with migraine and episodic ataxia
PT type 2 - useful for diagnosis and development of specific treatments
XX
PS Disclosure; Fig 1; 157pp; English.
XX
CC Sequences shown in V29330 to V29371 represent the 47 exons and flanking
CC intronic sequences containing the complete coding region of the human
CC calcium ion channel alpha 1 subunit gene and part of untranslated
CC sequences. The channel is related to familial hemiplegic migraine (FHM)
CC and/or episodic ataxia type 2 (EA-2) and is derived from, related to or
CC associated with a gene present in humans on chromosome 19p13.1-13.2. The
CC encoding gene can be used to localise or identify genes related to
CC episodic neurological disorders, specifically migraine, FHM or EA-2, but
CC also epilepsy. The isolated or a recombinant nucleic acid can also be
CC used to distinguish between alleles of the corresponding gene. Cells and
CC animals containing recombinant expression vectors comprising the nucleic
CC acid can be useful in study, development and treatment of migraine, FHM,
CC EA-2 and epilepsy. Proteins or peptides encoded by the nucleic acid and
CC natural or synthetic antibodies against the proteins can be used to
CC diagnose FHM, EA-2, migraine and other neurological conditions associated
CC with cation channel disfunction.
XX
SQ Sequence 559 BP; 119 A; 157 C; 160 G; 120 T; 3 other;

Query Match 0.6%; Score 17; DB 19; Length 559;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1207 aaacccagctcaacctc 1223
||||||||||||||||
Db 339 AAACCCAGCTCAACCTC 323

RESULT 34
X24995/c
ID X24995 standard; cDNA; 590 BP.

XX
XC X24995;
XX
XI 05-JUL-1999 (first entry)
XX
XE Murine Bcl-2 interacting mediator of cell death Bim-EL cDNA.
XX
XW Bim-EL; Bcl-2 interacting mediator of cell death; apoptosis;
XW cell cycle; mouse; cancer; autoimmune disease;
XW degenerative disease; therapy; contraceptive; splice variant;
XW isoform; ss.
XX
XS Mus musculus.
XX
XN WO9914321-A1.
XX
XD 25-MAR-1999.
XX
XF 17-SEP-1998; 98WO-AU00772.
XX
XR 24-SEP-1997; 97AU-0009373.
XR 17-SEP-1997; 97AU-0009263.
XX
XA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
XI Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
XI Puthalakath H, Strasser A;
XX
XR WPI; 1999-244030/20.
XR P-PSDB; W98156.
XX
XT New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
XT treatment
XX
XS Claim 3; Page 96-97; 145pp; English.
XX
XC The present sequence encodes the extra long form (EL) of murine Bim,
XC or Bcl-2 interacting mediator of cell death (see W98156), a novel
XC member of the Bcl-2 family that is capable of inducing cell death
XC (apoptosis) and which acts as a 'death-ligand' for certain members
XC of the pro-survival Bcl-2 family. Bim is a BH3-only protein, as the
XC only Bcl-2 homology region which it encompasses is BH3. It is the
XC only BH3-only protein for which splice variants exist. These
XC result in the expression of a variety of isoforms, i.e. Bim-S,
XC Bim-L and Bim-EL (see W98154-56). cDNAs encoding these murine Bim
XC isoforms were obtained from a T lymphoma cDNA library using human
XC recombinant Bcl-2 protein. The murine Bim gene has been mapped to
XC chromosome 2 at bands F3-G. Human Bim-L and Bim-EL isoforms have
XC also been identified (see W98157-58). Binding the dynein light
XC chain was shown to regulate the pro-apoptotic activity of Bim.
XC Bim-S, the splice variant which does not bind to dynein light
XC chain, is a much more potent killer than either Bim-L or Bim-EL.
XC The invention provides variants (see W98159-68) of murine and human
XC Bim-L or Bim-EL that cannot bind, couple or otherwise associate
XC with a dynein light chain. The identification of Bim permits the
XC identification and rational design of a range of products for use
XC in therapy, diagnosis, antibody generation and involving modulation
XC of physiological cell death. These therapeutic molecules may act
XC as either antagonists or agonists of Bim's function and will be
XC useful in cancer, autoimmune or degenerative disease therapy.
XC Increased Bim expression or Bim activity is useful, e.g. for
XC treatment or prophylaxis in conditions such as cancer and deletion
XC of autoreactive lymphocytes in autoimmune disease. Decreased Bim
XC expression of Bim activity is useful in regulating inhibition or
XC prevention of cell death or degeneration such as under cytotoxic
XC conditions during e.g. gamma-irradiation and chemotherapy or during
XC HIV/AIDS or other viral infections, ischemia, myocardial infarction,
XC hypoxia, degenerative diseases or for prolonging the survival of
XC cells being transplanted for treatment of disease. Since Bim is
XC expressed in germ cells, modulating Bim expression or Bim activity
XC is useful, e.g. as a contraceptive or method of sterilization by
XC preventing generation of fertile sperm.
XX

SQ Sequence 590 BP; 137 A; 178 C; 150 G; 125 T; 0 other;

Query Match 0.6%; Score 17; DB 20; Length 590;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 654 aggcctctcagcaggct 670
||||||||||||||||||
Db 75 AGGCCTCTCAGCAGGCT 59

RESULT 35
V87198/c
ID V87198 standard; cDNA; 605 BP.
XX
AC V87198;
XX
DT 27-APR-1999 (first entry)
XX
DE EST clone BN180.
XX
KW Expressed sequence tag; secreted protein; haematopoiesis regulator;
KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX
OS Homo sapiens.
XX
PN WO9845435-A2.
XX
PD 15-OCT-1998.
XX
PF 10-APR-1998; 98WO-US06954.
XX
PR 10-APR-1997; 97US-0835913.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
PI Racie LA, Spaulding V, Treacy M;
XX
DR WPI; 1999-070076/06.
XX
PT New polynucleotides encoding human secreted proteins - derived from
PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
PT ovary, pituitary, retina and colon cDNA libraries
XX
PS Claim 1; Page 486; 633pp; English.
XX
CC This sequence represents an expressed sequence tag (EST), and is a
CC polynucleotide of the invention. The polynucleotides of the invention are
CC all secreted EST sequences isolated from a variety of human tissue
CC sources. The EST sequences and proteins encoded by them are predicted to
CC have useful biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested activities
CC include nutritional activity, immune stimulating or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
CC activity. The EST sequences are also stated to be useful for gene
CC therapy.
XX
SQ Sequence 605 BP; 149 A; 126 C; 194 G; 136 T; 0 other;

Query Match 0.6%; Score 17; DB 20; Length 605;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 gttttggtggagacggg 47

```
          ||||||||||||||||||
Db    212 GTTTTGGTGGAGACGGG 196


RESULT  36
Z24979/c
ID   Z24979 standard; cDNA; 791 BP.
XX
AC   Z24979;
XX
DT   05-JUL-1999  (first entry)
XX
DE   Clone GJ156 encoding TRAIN-R secreted form C-terminus.
XX
KW   TRAIN-R; receptor; human; tumour necrosis factor receptor;
KW   agonist; antagonist; cancer; immunological disease; therapy;
KW   cytostatic; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   intron          1..350
FT                   /*tag= a
FT   exon            351..790
FT                   /*tag= b
FT   CDS             352..444
FT                   /*tag= c
FT                   /partial
FT                   /product= "TRAIN-R secreted form C-terminus"
FT   3'UTR           45..790
FT                   /*tag= d
XX
PN   WO9913078-A1.
XX
PD   18-MAR-1999.
XX
PF   11-SEP-1998;   98WO-US19030.
XX
PR   06-MAY-1998;   98US-0084422.
PR   12-SEP-1997;   97US-0058631.
XX
PA   (BIOJ ) BIOGEN INC.
XX
PI   Hession C,  Tschopp J;
XX
DR   WPI; 1999-229238/19.
DR   P-PSDB; W98147.
XX
PT   New cysteine-rich tumor necrosis factor receptor
XX
PS   Claim 1; Page 28; 30pp; English.
XX
CC   The present sequence includes an exon encoding the C-terminus (see
CC   W98147) of a soluble form of a novel human cysteine-rich tumour
CC   necrosis factor receptor family member termed TRAIN-R.  It comprises
CC   clone GJ156, obtained from a Clontech human adult lung cDNA library.
CC   The encoded 30-amino acid C-terminal peptide is identical to amino
CC   acids 121-149 of the composite TRAIN-R protein given in W98146 and
CC   to amino acids 121-150 of the C-terminus of murine TRAIN-R short
CC   form (secreted protein, see W98144).  The soluble protein is
CC   expected to inhibit signalling by the full-length TRAIN-R.  Human
CC   TRAIN-R is expressed at low levels in every tissue and cell line
CC   tested thus far, with higher expression detected in heart, prostate,
CC   ovary, testis, peripheral blood lymphocytes, thyroid and adrenal
CC   gland.  Cell death can be induced by administering an agent capable
CC   of inhibiting the binding of TRAIN-R to its ligand.  A claimed method
CC   of treating, or reducing, the advancement, severity or effects of
CC   an immunological disease in a mammal comprises administering a
CC   pharmaceutical composition which comprises a TRAIN-R blocking agent,
CC   e.g. soluble TRAIN-R.  TRAIN-R can be fused to an immunoglobulin to
CC   produce a fusion protein which may be targeted to various sites.
CC   It can be used in binding assays, and to identify antagonists and
CC   agonists.  Anti-TRAIN-R antibodies can be used to reduce the
CC   severity of an immune response or to treat cancer.  TRAIN-R
CC   blocking agents can also be used to reduce the severity or effects
CC   of an immunological disease (all claimed).
XX
SQ   Sequence 791 BP; 202 A; 189 C; 165 G; 235 T; 0 other;


  Query Match              0.6%;  Score 17;  DB 20;  Length 791;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2130 gaagaggaagcagtgga 2146
          |||||||||||||||||
Db    203 GAAGAGGAAGCAGTGGA 187


RESULT  37
Z91823
ID   Z91823 standard; DNA; 855 BP.
XX
AC   Z91823;
XX
DT   02-JUN-2000  (first entry)
XX
DE   Streptococcus pneumoniae DNA sequence ID33.
XX
KW   Streptococcus pneumoniae infection; immunogen; antigen; diagnosis; AIDS;
KW   bacterial pneumonia; asplenia; heart disease; lung disease; alcoholism;
KW   kidney disease; diabetes; immunosuppressive disorder; otitis media;
KW   pneumococcal septicaemia; sinusitis; meningitis; therapy; ss.
XX
OS   Streptococcus pneumoniae.
XX
PN   WO200006738-A2.
XX
PD   10-FEB-2000.
XX
PF   27-JUL-1999;   99WO-GB02452.
XX
PR   27-JUL-1998;   98GB-0016336.
PR   19-MAR-1999;   99US-0125329.
XX
PA   (MICR-) MICROBIAL TECHNICS LTD.
XX
PI   Le Page RWF,  Wells JM,  Hanniffy SB,  Hansbro PM;
XX
DR   WPI; 2000-195301/17.
DR   P-PSDB; Y81727.
XX
PT   Streptococcal proteins and polynucleotides useful for diagnosis,
PT   treatment and prophylaxis of bacterial infections
XX
PS   Claim 2; Page 47-48; 76pp; English.
XX
CC   This sequence encodes a Streptococcus pneumoniae protein of the
CC   invention. The proteins (or their homologues, derivatives and/or
CC   fragments)are useful as immunogens or antigens. Immunogenic or antigenic
CC   compositions comprising the proteins are useful as vaccines and also in
CC   diagnostic assays. The sequences are useful for the detection or
CC   diagnosis of S. pneumoniae infection, by contacting a sample to be tested
CC   with them. Agents capable of antagonising, inhibiting or interfering with
CC   the function or expression of the protein or polypeptide are useful in
CC   medical compositions in the treatment or prophylaxis of S. pneumoniae
CC   infection. As the sequences can be used to treat S. pneumoniae infection,
CC   they can be used to treat bacterial pneumonia, which has high rates in
CC   young children, the elderly, and in patients with predisposing conditions
CC   such as asplenia, heart, lung and kidney disease, diabetes, alcoholism,
CC   or with immunosuppressive disorders, especially AIDS. They can also be
CC   used to treat pneumococcal septicaemia, otitis media, sinusitis, and
CC   meningitis.
XX
SQ   Sequence 855 BP; 235 A; 173 C; 202 G; 245 T; 0 other;
```

```
  Query Match             0.6%;  Score 17;  DB 21;  Length 855;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1947 gatttggaagagtttca 1963
          |||||||||||||||||
Db    139 gatttggaagagtttca 155


RESULT  38
X40199
ID   X40199 standard; DNA; 1022 BP.
XX
AC   X40199;
XX
DT   02-JUL-1999  (first entry)
XX
DE   MAGE-4 encoding gene.
XX
KW   Cancer associated antigen; diagnosis; research; treatment; human;
KW   breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW   prostate cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO9904265-A2.
XX
PD   28-JAN-1999.
XX
PF   15-JUL-1998;   98WO-US14679.
XX
PR   22-JUN-1998;   98US-0102322.
PR   17-JUL-1997;   97US-0896164.
PR   10-OCT-1997;   97US-0061599.
PR   10-OCT-1997;   97US-0061765.
PR   10-OCT-1997;   97US-0948705.
PR   11-OCT-1997;   97GB-0021697.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
XX
PI   Chen Y,  Gout I,  Gure A,  O'Hare M,  Obata Y,  Old LJ;
PI   Pfreundschuh M,  Sahin U,  Scanlan MJ,  Stockert E;
PI   Tureci O;
XX
DR   WPI; 1999-132448/11.
DR   P-PSDB; Y06998.
XX
PT   New isolated cancer associated nucleic acids and polypeptides -
PT   isolated using sera from cancer patients, used to develop products
PT   for the diagnosis, monitoring or treatment of cancers
XX
PS   Claim 67; Page 780; 787pp; English.
XX
CC   The invention relates to a method for diagnosing a disorder characterised
CC   by expression of a human cancer associated antigen precursor coded for by
CC   a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC   biological sample isolated from a subject with an agent that specifically
CC   binds to the NAM, an expression product or a fragment of an expression
CC   product complexed with an HLA molecule; and (b) determining the
CC   interaction between the agent and the NAM or the expression product as a
CC   determination of the disorder. The products and methods can be used in
CC   the diagnosis, monitoring, research, or treatment of conditions
CC   characterised by the expression of various cancer associated antigens.
CC   The invention provides nucleic acid sequences and encoded polypeptides
CC   which are cancer associated antigen precursors expressed in human breast
CC   cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC   lung cancer.
XX
SQ   Sequence 1022 BP; 230 A; 273 C; 302 G; 217 T; 0 other;

  Query Match             0.6%;  Score 17;  DB 20;  Length 1022;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1652 ggtcctgggcaccctgg 1668
          |||||||||||||||||
Db    203 ggtcctgggcaccctgg 219


RESULT  39
Q67866
ID   Q67866 standard; DNA; 1084 BP.
XX
AC   Q67866;
XX
DT   22-MAR-1995  (first entry)
XX
DE   H6/MAGE-1 expression cassette from pMAW037.
XX
KW   Polymerase chain reaction; primer; amplify; NYVAC; ALVAC; recombinant;
KW   human; MAGE-1; melanoma-associated antigen; MZ2-E; testis; pTZ18RMAGE1;
KW   primary melanoma tumour cell; melanoma-derived cell line; tumour;
KW   poxvirus; antigenic response; immunological response; pathogen; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1..51
FT                   /*tag= a
FT                   /note= "Flanking sequence"
FT   promoter        52..178
FT                   /*tag= b
FT                   /note= "Vaccinia H6 promoter"
FT   CDS             179..1009
FT                   /*tag= c
FT                   /product= MAGE-1
FT   misc_feature    1010..1084
FT                   /*tag= d
FT                   /note= "Flanking sequence"
XX
PN   WO9416716-A.
XX
PD   04-AUG-1994.
XX
PF   21-JAN-1994;   94WO-US00888.
XX
PR   21-JAN-1993;   93US-0007115.
PR   19-JAN-1994;   94US-0184009.
XX
PA   (VIRO-) VIROGENETICS CORP.
XX
PI   Cox WI,  Paoletti E,  Tartaglia J;
XX
DR   WPI; 1994-263767/32.
XX
PT   Attenuated recombinant virus used for cancer therapy - comprises
PT   DNA encoding cytokine and/or tumour associated antigen
XX
PS   Example 16; Fig 20; 232pp; English.
XX
CC   The sequences given in Q67865-66 represent expression cassettes
CC   containing the vaccinia H6 promoter and the human MAGE-1 gene which
CC   encodes human melanoma-associated antigen MZ2-E, in vCP235 and pMAW037,
CC   respectively.  These sequences were used in the construction of NYVAC-
CC   and ALVAC-based recombinant viruses containing the MAGE-1 gene.  MAGE-1
CC   is expressed in primary melanoma tumour cells, melanoma-derived cell
CC   lines and certian tumours of non-melanoma origins but not in normal
CC   cells except in testis.  A first PCR fragment containing the last 18 bp
CC   and the initial 24 nucleotides of the MAGE-1 gene was generated and
CC   ligated to a second PCR fragment amplified from plasmid pTZ18RMAGE1
CC   which contains the initial 546 bp of the MAGE-1 coding sequence.  The
CC   terminal sequence of MAGE-1 was amplified and a fusion product was
CC   generated containing the H6 promoter and the full length MAGE-1 sequence.
CC   This construct may be introduced in to the poxvirus derived plasmids,
```

CC ALVAC and NYVAC. The resulting viruses may be used in a composition for
CC inducing an antigenic or immunological response, ie. for immunisation
CC against pathogens.
XX
SQ Sequence 1084 BP; 266 A; 256 C; 280 G; 282 T; 0 other;

Query Match 0.6%; Score 17; DB 15; Length 1084;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1652 ggtcctgggcaccctgg 1668
||||||||||||||||||
Db 292 ggtcctgggcaccctgg 308

RESULT 40
Z08442
ID Z08442 standard; DNA; 1084 BP.
XX
AC Z08442;
XX
DT 19-OCT-1999 (first entry)
XX
DE H6/MAGE-1 expression cassette and flanking regions from pMAW037.
XX
KW Attenuated recombinant virus; cytokine; tumour associated antigen;
KW NYVAC recombinant virus; ALVAC recombinant virus; gene therapy; rabies;
KW cancer; tumour necrosis factor; nuclear phosphoprotein; p53; IL-2; GMCSF;
KW interleukin; interferon; IFN-gamma; IL-4; melanoma associated antigen;
KW carcinoembryonic antigen; immunisation; antigenic; poxvirus; influenza;
KW immunological response; immunotherapy; vaccine; Newcastle Disease; ss.
XX
OS Synthetic.
OS Homo sapiens.
OS Vaccinia virus.
XX
PN US5942235-A.
XX
PD 24-AUG-1999.
XX
PF 02-JUN-1995; 95US-0458356.
XX
PR 02-JUN-1995; 95US-0458356.
PR 24-DEC-1981; 81US-0334456.
PR 08-DEC-1982; 82US-0446824.
PR 19-JUN-1984; 84US-0622135.
PR 27-AUG-1987; 87US-0090209.
PR 28-AUG-1987; 87US-0090711.
PR 20-OCT-1987; 87US-0110335.
PR 25-APR-1988; 88US-0186054.
PR 23-AUG-1988; 88US-0234390.
PR 08-MAR-1989; 89US-0320471.
PR 14-FEB-1990; 90US-0478179.
PR 14-JUN-1990; 90US-0537882.
PR 14-JUN-1990; 90US-0537890.
PR 07-JAN-1991; 91US-0638080.
PR 07-MAR-1991; 91US-0666056.
PR 11-JUN-1991; 91US-0713967.
PR 16-DEC-1991; 91US-0805567.
PR 03-MAR-1992; 92US-0847977.
PR 06-MAR-1992; 92US-0847951.
PR 04-MAY-1992; 92US-0881995.
PR 22-JUL-1992; 92US-0918278.
PR 20-JAN-1993; 93US-0007115.
PR 19-JAN-1994; 94US-0184009.
PR 14-APR-1994; 94US-0228926.
PR 13-SEP-1994; 94US-0306259.
XX
PA (HEAL-) HEALTH RES INC.
XX
PI Paoletti E;
XX
DR WPI; 1999-493494/41.
XX
PT Recombinant poxviruses comprising exogenous DNA encoding antigenic
PT determinants useful in immunotherapy to immunize against cancers and
PT other diseases such as influenza, Newcastle Disease and rabies
XX
PS Example 16; Fig 20; 163pp; English.
XX
CC The present invention describes a recombinant poxvirus (I), comprising
CC exogenous DNA encoding an antigenic determinant of a pathogen which is
CC then expressed in vivo in infected host cells after administration to a
CC patient and therefore induces an immunological response. (I) may be used
CC to vaccinate patients against a wide range of diseases and disorders
CC depending on the type of antigen encoded by the exogenous DNA. (I) may
CC be used to vaccinate against diseases such as rabies, influenza and
CC Newcastle Disease. It is particularly useful for immunising against
CC cancers. The poxvirus (I) also provides a means of manipulating
CC lymphocytes and tumour cells for use in cell-based immunotherapeutic
CC modalities for cancer. (I) also have enhanced safety compared to
CC unattenuated viruses (attenuation reduces the virulence of the viruses)
CC and known recombinant poxvirus vaccines. This increased level of safety
CC reduces the possibility of a 'runaway' infection in the host and reduces
CC the chance of transmission from vaccinated to unvaccinated individuals
CC and contamination of the environment. The present sequence represents a
CC H6/MAGE-1 expression cassette and flanking regions from pMAW037
CC used in the exemplification of the present invention.
XX
SQ Sequence 1084 BP; 266 A; 256 C; 280 G; 282 T; 0 other;

Query Match 0.6%; Score 17; DB 20; Length 1084;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1652 ggtcctgggcaccctgg 1668
||||||||||||||||||
Db 292 ggtcctgggcaccctgg 308

RESULT 41
Q67865
ID Q67865 standard; DNA; 1094 BP.
XX
AC Q67865;
XX
DT 22-MAR-1995 (first entry)
XX
DE H6/MAGE-1 expression cassette from vCP235.
XX
KW Polymerase chain reaction; primer; amplify; NYVAC; ALVAC; recombinant;
KW human; MAGE-1; melanoma-associated antigen; MZ2-E; testis; pTZ18RMAGE1;
KW primary melanoma tumour cell; melanoma-derived cell line; tumour;
KW poxvirus; antigenic response; immunological response; pathogen; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT promoter 74..200
FT /*tag= a
FT /note= "Vaccinia H6 promoter"
FT CDS 201..1031
FT /*tag= b
FT /product= MAGE-1
FT misc_feature 1032..1094
FT /*tag= c
FT /note= "Flanking sequence"
XX
PN WO9416716-A.
XX
PD 04-AUG-1994.
XX
PF 21-JAN-1994; 94WO-US00888.
XX

```
PR   21-JAN-1993;   93US-0007115.
PR   19-JAN-1994;   94US-0184009.
XX
PA   (VIRO-) VIROGENETICS CORP.
XX
PI   Cox WI,  Paoletti E,  Tartaglia J;
XX
DR   WPI; 1994-263767/32.
XX
PT   Attenuated recombinant virus used for cancer therapy - comprises
PT   DNA encoding cytokine and/or tumour associated antigen
XX
PS   Example 16; Fig 19; 232pp; English.
XX
CC   The sequences given in Q67865-66 represent expression cassettes
CC   containing the vaccinia H6 promoter and the human MAGE-1 gene which
CC   encodes human melanoma-associated antigen MZ2-E, in vCP235 and pMAW037,
CC   respectively.  These sequences were used in the construction of NYVAC-
CC   and ALVAC-based recombinant viruses containing the MAGE-1 gene.  MAGE-1
CC   is expressed in primary melanoma tumour cells, melanoma-derived cell
CC   lines and certian tumours of non-melanoma origins but not in normal
CC   cells except in testis.  A first PCR fragment containing the last 18 bp
CC   and the initial 24 nucleotides of the MAGE-1 gene was generated and
CC   ligated to a second PCR fragment amplified from plasmid pTZ18RMAGE1
CC   which contains the initial 546 bp of the MAGE-1 coding sequence.  The
CC   terminal sequence of MAGE-1 was amplified and a fusion product was
CC   generated containing the H6 promoter and the full length MAGE-1 sequence.
CC   This construct may be introduced in to the poxvirus derived plasmids,
CC   ALVAC and NYVAC.  The resulting viruses may be used in a composition for
CC   inducing an antigenic or immunological response, ie. for immunisation
CC   against pathogens.
XX
SQ   Sequence 1094 BP; 269 A; 259 C; 293 G; 273 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 15;  Length 1094;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1652 ggtcctgggcaccctgg 1668
           |||||||||||||||||
Db    314 ggtcctgggcaccctgg 330


RESULT  42
Z08441
ID   Z08441 standard; DNA; 1094 BP.
XX
AC   Z08441;
XX
DT   19-OCT-1999  (first entry)
XX
DE   H6/MAGE-1 expression cassette and flanking regions from vCP235.
XX
KW   Attenuated recombinant virus; cytokine; tumour associated antigen;
KW   NYVAC recombinant virus; ALVAC recombinant virus; gene therapy; rabies;
KW   cancer; tumour necrosis factor; nuclear phosphoprotein; p53; IL-2; GMCSF;
KW   interleukin; interferon; IFN-gamma; IL-4; melanoma associated antigen;
KW   carcinoembryonic antigen; immunisation; antigenic; poxvirus; influenza;
KW   immunological response; immunotherapy; vaccine; Newcastle Disease; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
OS   Vaccinia virus.
XX
PN   US5942235-A.
XX
PD   24-AUG-1999.
XX
PF   02-JUN-1995;   95US-0458356.
XX
PR   02-JUN-1995;   95US-0458356.
PR   24-DEC-1981;   81US-0334456.
PR   08-DEC-1982;   82US-0446824.
PR   19-JUN-1984;   84US-0622135.
PR   27-AUG-1987;   87US-0090209.
PR   28-AUG-1987;   87US-0090711.
PR   20-OCT-1987;   87US-0110335.
PR   25-APR-1988;   88US-0186054.
PR   23-AUG-1988;   88US-0234390.
PR   08-MAR-1989;   89US-0320471.
PR   14-FEB-1990;   90US-0478179.
PR   14-JUN-1990;   90US-0537882.
PR   14-JUN-1990;   90US-0537890.
PR   07-JAN-1991;   91US-0638080.
PR   07-MAR-1991;   91US-0666056.
PR   11-JUN-1991;   91US-0713967.
PR   16-DEC-1991;   91US-0805567.
PR   03-MAR-1992;   92US-0847977.
PR   06-MAR-1992;   92US-0847951.
PR   04-MAY-1992;   92US-0881995.
PR   22-JUL-1992;   92US-0918278.
PR   20-JAN-1993;   93US-0007115.
PR   19-JAN-1994;   94US-0184009.
PR   14-APR-1994;   94US-0228926.
PR   13-SEP-1994;   94US-0306259.
XX
PA   (HEAL-) HEALTH RES INC.
XX
PI   Paoletti E;
XX
DR   WPI; 1999-493494/41.
XX
PT   Recombinant poxviruses comprising exogenous DNA encoding antigenic
PT   determinants useful in immunotherapy to immunize against cancers and
PT   other diseases such as influenza, Newcastle Disease and rabies
XX
PS   Example 16; Fig 19; 163pp; English.
XX
CC   The present invention describes a recombinant poxvirus (I), comprising
CC   exogenous DNA encoding an antigenic determinant of a pathogen which is
CC   then expressed in vivo in infected host cells after administration to a
CC   patient and therefore induces an immunological response. (I) may be used
CC   to vaccinate patients against a wide range of diseases and disorders
CC   depending on the type of antigen encoded by the exogenous DNA. (I) may
CC   be used to vaccinate against diseases such as rabies, influenza and
CC   Newcastle Disease. It is particularly useful for immunising against
CC   cancers. The poxvirus (I) also provides a means of manipulating
CC   lymphocytes and tumour cells for use in cell-based immunotherapeutic
CC   modalities for cancer. (I) also have enhanced safety compared to
CC   unattenuated viruses (attenuation reduces the virulence of the viruses)
CC   and known recombinant poxvirus vaccines. This increased level of safety
CC   reduces the possibility of a 'runaway' infection in the host and reduces
CC   the chance of transmission from vaccinated to unvaccinated individuals
CC   and contamination of the environment. The present sequence represents a
CC   H6/MAGE-1 expression cassette and flanking regions from vCP235
CC   used in the exemplification of the present invention.
XX
SQ   Sequence 1094 BP; 269 A; 259 C; 293 G; 273 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 20;  Length 1094;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1652 ggtcctgggcaccctgg 1668
           |||||||||||||||||
Db    314 ggtcctgggcaccctgg 330


RESULT  43
X27353
ID   X27353 standard; DNA; 1140 BP.
XX
AC   X27353;
XX
```

```
RT   11-JUN-1999  (first entry)
XX
DE   Human secreted protein gene 43 clone HTADX17.
XX
KW   Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW   diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW   developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW   immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW   inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW   cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW   osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW   endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS   Homo sapiens.
XX
PN   WO9902546-A1.
XX
PD   21-JAN-1999.
XX
PF   07-JUL-1998;   98WO-US13684.
XX
PR   12-SEP-1997;   97US-0058785.
PR   08-JUL-1997;   97US-0051916.
PR   08-JUL-1997;   97US-0051918.
PR   08-JUL-1997;   97US-0051919.
PR   08-JUL-1997;   97US-0051920.
PR   08-JUL-1997;   97US-0051925.
PR   08-JUL-1997;   97US-0051926.
PR   08-JUL-1997;   97US-0051928.
PR   08-JUL-1997;   97US-0051929.
PR   08-JUL-1997;   97US-0051930.
PR   08-JUL-1997;   97US-0051931.
PR   08-JUL-1997;   97US-0051932.
PR   08-JUL-1997;   97US-0052732.
PR   08-JUL-1997;   97US-0052733.
PR   08-JUL-1997;   97US-0052793.
PR   08-JUL-1997;   97US-0052795.
PR   08-JUL-1997;   97US-0052803.
PR   18-AUG-1997;   97US-0055684.
PR   18-AUG-1997;   97US-0055722.
PR   18-AUG-1997;   97US-0055723.
PR   18-AUG-1997;   97US-0055947.
PR   18-AUG-1997;   97US-0055948.
PR   18-AUG-1997;   97US-0055949.
PR   18-AUG-1997;   97US-0055950.
PR   18-AUG-1997;   97US-0055953.
PR   18-AUG-1997;   97US-0055954.
PR   18-AUG-1997;   97US-0055964.
PR   18-AUG-1997;   97US-0055984.
PR   18-AUG-1997;   97US-0056360.
PR   12-SEP-1997;   97US-0058660.
PR   12-SEP-1997;   97US-0058661.
PR   12-SEP-1997;   97US-0058664.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Brewer LA,  Ebner R,  Fischer CL,  Kyaw H,  Lafleur DW,  Li Y,  Moore PA;
PI   Olsen HS,  Rosen CA,  Ruben SM,  Shi Y,  Soppet DR,  Zeng Z;
XX
DR   WPI; 1999-120770/10.
DR   P-PSDB; Y02692.
XX
PT   New isolated human genes and the secreted polypeptides they encode -
PT   useful for diagnosis and treatment of e.g. cancers, neurological
PT   disorders, immune diseases, inflammation or blood disorders
XX
PS   Claim 1; Page 271; 464pp; English.
XX
CC   This sequence represents a nucleic acid molecule which encodes a secreted
CC   human protein.  The gene number, and the clone it is derived from, are
CC   detailed in the descriptor line.  The gene can be used to generate fusion
CC   proteins by linking to the gene to a human immunoglobulin Fc portion
CC   (e.g. X27302) for increasing the stability of the fused protein as
CC   compared to the human protein only.
CC   The invention relates to 123 novel genes and their fragments (nucleic
CC   acid sequences: X27311-X27449; amino acid sequences Y02650-Y02788) which
CC   are useful for preventing, treating or ameliorating  medical conditions
CC   e.g. by protein or gene therapy.  Also, pathological conditions can be
CC   diagnosed by determining the amount of the new polypeptides in a sample
CC   or by determining the presence of mutations in the new polynucleotides.
CC   Specific uses are described for each of the 123 polynucleotides, based on
CC   which tissues they are most highly expressed in (see X27311 for described
CC   uses).
XX
SQ   Sequence 1140 BP; 280 A; 312 C; 289 G; 254 T; 5 other;


  Query Match             0.6%;  Score 17;  DB 20;  Length 1140;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1802 caaagcctggctccagc 1818
           |||||||||||||||||
Db    930 caaagcctggctccagc 946


RESULT  44
Q29634
ID   Q29634 standard; DNA; 1173 BP.
XX
AC   Q29634;
XX
DT   16-MAR-1993  (first entry)
XX
DE   Hepatitis C virus HC-J5 3' region.
XX
KW   Non-A non-B hepatitis; NANBH; HCV; detection; diagnosis; screening;
KW   PCR; primer; polymerase chain reaction; ss.
XX
OS   Hepatitis C virus.
XX
PN   EP510952-A.
XX
PD   28-OCT-1992.
XX
PF   23-APR-1992;   92EP-0303625.
XX
PR   26-APR-1991;   91JP-0191376.
XX
PA   (IMMO ) IMMUNO JAPAN INC.
XX
PI   Nakamura T,  Okamoto H;
XX
DR   WPI; 1992-359137/44.
XX
PT   Detection of non-A, non-B hepatitis virus - using new
PT   oligo-nucleotide primers with nucleotide sequences corresp. to
PT   part. of the viral RNA
XX
PS   Disclosure; Page 28; 54pp; English.
XX
CC   This sequence represents the 3' region of hepatitis C virus RNA. The
CC   original sample was obtained from human and chimpanzee plasma. RNA
CC   was isolated from several samples and homology compared, and the
CC   respective sequence of about 1900 - 2500 nucleotides of the 5'
CC   terminus and 1100 nucleotides of the 3' terminus determined.
XX
SQ   Sequence 1173 BP; 246 A; 358 C; 310 G; 259 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 13;  Length 1173;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    294 cggtatctcttcaactg 310
          |||||||||||||||||
```

```
Db    910 cggtatctcttcaactg 926


RESULT  45
Q29635
ID   Q29635 standard; DNA; 1173 BP.
XX
AC   Q29635;
XX
DT   16-MAR-1993  (first entry)
XX
DE   Hepatitis C virus HC-J6 3' region.
XX
KW   Non-A non-B hepatitis; NANBH; HCV; detection; diagnosis; screening;
KW   PCR; primer; polymerase chain reaction; ss.
XX
OS   Hepatitis C virus.
XX
PN   EP510952-A.
XX
PD   28-OCT-1992.
XX
PF   23-APR-1992;   92EP-0303625.
XX
PR   26-APR-1991;   91JP-0191376.
XX
PA   (IMMO ) IMMUNO JAPAN INC.
XX
PI   Nakamura T,  Okamoto H;
XX
DR   WPI; 1992-359137/44.
XX
PT   Detection of non-A, non-B hepatitis virus - using new
PT   oligo-nucleotide primers with nucleotide sequences corresp. to
PT   part. of the viral RNA
XX
PS   Disclosure; Page 29; 54pp; English.
XX
CC   This sequence represents the 3' region of hepatitis C virus RNA. The
CC   original sample was obtained from human and chimpanzee plasma. RNA
CC   was isolated from several samples and homology compared, and the
CC   respective sequence of about 1900 - 2500 nucleotides of the 5'
CC   terminus and 1100 nucleotides of the 3' terminus determined.
XX
SQ   Sequence 1173 BP; 251 A; 362 C; 304 G; 256 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 13;  Length 1173;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    294 cggtatctcttcaactg 310
          |||||||||||||||||
Db    910 cggtatctcttcaactg 926


RESULT  46
Q43895
ID   Q43895 standard; cDNA to RNA; 1173 BP.
XX
AC   Q43895;
XX
DT   21-OCT-1993  (first entry)
XX
DE   NANB hepatitis virus polynucleotide N-1173-3.
XX
KW   Non-A, non-B; virus; polymerase chain reaction; detection;
KW   sensitive; specific; HCV; NANBH; ss.
XX
OS   Non-A, non-B hepatitis virus.
XX
FH   Key             Location/Qualifiers
FT   CDS             2..1119
FT                   /*tag=  a
XX
PN   JP05091884-A.
XX
PD   16-APR-1993.
XX
PF   10-APR-1991;   91JP-0196175.
XX
PR   12-JUN-1990;   90JP-0153401.
PR   08-NOV-1990;   90JP-0304405.
XX
PA   (NAKA/) NAKAMURA T.
XX
DR   WPI; 1993-199637/25.
DR   P-PSDB; R38285.
XX
PT   Antigen related to non-A and non-B hepatitis virus - comprises
PT   non-translation region comprising 340 - 341 mols. of nucleotides,
PT   non-translation region comprising 1885 - 2551 mols. of
PT   nucleotides including region 1,149 and, etc.
XX
PS   Claim 9; Page 29; 73pp; Japanese.
XX
CC   The sequence is that of NANB hepatitis virus polynucleotide N-1173-3
CC   which codes for a non-A, non-B (NANB) hepatitis virus gene HC-OM.
CC   The polypeptide it encodes may be used in a system for detecting
CC   NANB hepatitis. This method is highly specific and sensitive, and
CC   can detect NANB hepatitis virus which could not be detected by
CC   conventional methods.
XX
SQ   Sequence 1173 BP; 247 A; 358 C; 309 G; 259 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 14;  Length 1173;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    294 cggtatctcttcaactg 310
          |||||||||||||||||
Db    910 cggtatctcttcaactg 926


RESULT  47
Z46089
ID   Z46089 standard; cDNA; 1209 BP.
XX
AC   Z46089;
XX
DT   05-MAY-2000  (first entry)
XX
DE   cDNA encoding a forkhead activin signal transducer designated FAST2.
XX
KW   Forkhead activin signal transducer protein; FAST2; activin signalling;
KW   winged-helix/forkhead domain protein; homeobox gene; goosecoid inducer;
KW   gsc; transforming growth factor-beta signalling; gsc promoter;
KW   signal transduction; transcription factor; wound healing; inflammation;
KW   tumour progression; scarring; arthritis; fibrosis; liver fibrosis;
KW   kidney fibrosis; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             4..1209
FT                   /*tag=  a
FT                   /product= "Forkhead activin signal transducer protein"
FT                   /transl_except= (pos: 1183..1185, aa: Cys)
XX
PN   WO200004143-A1.
XX
PD   27-JAN-2000.
XX
PF   19-JUL-1999;   99WO-CA00645.
XX
```

R 17-JUL-1998; 98CA-2237788.
X
A (HSCR-) HSC RES & DEV LP.
X
I Wrana JL, Attisano L;
X
R WPI; 2000-171267/15.
R P-PSDB; Y54601.
X
T New mammalian transcription factor, useful for preventing or treating
T disorders associated with transforming growth factor beta or activin
T signaling pathways -
X
S Claim 3; Page 43; 76pp; English.
X
C The present sequence encodes a mammalian forkhead activin signal
C transducer (FAST) protein, designated FAST2. The protein is a
C winged-helix/forkhead domain protein. The protein is an inducer of
C the homeobox gene goosecoid (gsc) by transforming growth factor
C (TGF)-beta or activin signaling. FAST2 binds to a nucleotide sequence
C in the gsc promoter. The FAST2 protein is useful for modulating signal
C transduction in a TGF-beta or activin signaling pathway, which involves
C FAST2 as transcription factor, by modulating the formation of
C Smad2/Smad4/FAST2 or Smad3/Smad4/FAST2 complex. Inhibition of FAST2
C binding to its target DNA site inhibits FAST2 specific TGF-beta
C signaling, which is associated with wound healing inflammation, and
C tumour progression. Excessive signaling is associated with scarring,
C arthritis and fibrosis in numerous diseases, including fibrosis of the
C liver and kidney.
X
Q Sequence 1209 BP; 247 A; 412 C; 315 G; 235 T; 0 other;

Query Match 0.6%; Score 17; DB 21; Length 1209;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1074 cacatggccccagcatc 1090
||||||||||||||||
Db 636 cacatggccccagcatc 652

ESULT 48
37593
D X87593 standard; cDNA; 1338 BP.
X
A X87593;
X
T 26-OCT-1999 (first entry)
X
D CLYTA-MAGE-1-His fusion DNA.
X
W MAGE-1; CLYTA-MAGE-1-His; fusion protein; tumour; melanoma;
W breast cancer; bladder cancer; lung cancer; colon cancer;
W head and squamous cell carcinoma; oesophagus carcinoma; vaccine;
W human; ss.
X
S Chimeric - Streptococcus pneumoniae.
S Chimeric - Homo sapiens.
S Chimeric - synthetic.
X
N WO9940188-A2.
X
D 12-AUG-1999.
X
F 02-FEB-1999; 99WO-EP00660.
X
R 06-FEB-1998; 98GB-0002650.
R 05-FEB-1998; 98GB-0002543.
X
A (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
X
I Cabezon Silva T, Cohen J, Slaoui MM, Vinals Bassols C;

XX
DR WPI; 1999-494293/41.
DR P-PSDB; Y06592.
XX
PT New protein derivatives used in cancer vaccine therapy for treating
PT a range of cancers including melanomas, carcinomas and cancers of
PT breast
XX
PS Example 9; Page 70-71; 72pp; English.
XX
CC This DNA sequence codes for a fusion protein (see Y06592) composed
CC of the C-terminal portion of the Streptococcus pneumoniae LYTA
CC protein (CLYTA), the human MAGE-1 tumour-associated antigen and a
CC hexahistidine tail. A vector designed for recombinant expression
CC of the fusion protein in Escherichia coli is provided. The CLYTA
CC moiety provides expression of soluble fusion protein, facilitates
CC affinity purification of the fusion protein, and also acts as a
CC T-helper epitope. The invention relates to MAGE proteins fused to
CC an immunological fusion partner, e.g. Clyta-MAGE-1-His. These novel
CC fusion proteins provide vaccines for immunotherapy of melanomas or
CC other MAGE-associated tumours like breast, bladder, lung and
CC non-small cell lung cancer, head and squamous cell carcinoma, colon
CC carcinoma and oesophagus carcinoma.
XX
SQ Sequence 1338 BP; 335 A; 334 C; 378 G; 291 T; 0 other;

Query Match 0.6%; Score 17; DB 20; Length 1338;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1652 ggtcctgggcaccctgg 1668
||||||||||||||||
Db 495 ggtcctgggcaccctgg 511

RESULT 49
X87591
ID X87591 standard; cDNA; 1341 BP.
XX
AC X87591;
XX
DT 26-OCT-1999 (first entry)
XX
DE Lipoprotein D-MAGE-1-His fusion DNA.
XX
KW MAGE-1; lipoprotein D; LPD-MAGE-1-His; fusion protein; tumour;
KW melanoma; breast cancer; bladder cancer; lung cancer;
KW head and squamous cell carcinoma; colon cancer;
KW oesophagus carcinoma; vaccine; human; ss.
XX
OS Chimeric - Haemophilus influenzae.
OS Chimeric - Homo sapiens.
OS Chimeric - synthetic.
XX
PN WO9940188-A2.
XX
PD 12-AUG-1999.
XX
PF 02-FEB-1999; 99WO-EP00660.
XX
PR 06-FEB-1998; 98GB-0002650.
PR 05-FEB-1998; 98GB-0002543.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Cabezon Silva T, Cohen J, Slaoui MM, Vinals Bassols C;
XX
DR WPI; 1999-494293/41.
DR P-PSDB; Y06590.
XX
PT New protein derivatives used in cancer vaccine therapy for treating
PT a range of cancers including melanomas, carcinomas and cancers of

```
DT   breast
XX
DS   Example 6; Page 66; 72pp; English.
XX
CC   This DNA sequence codes for a fusion protein (see Y06590) composed
CC   of lipidated protein D (LPD) of Haemophilus influenzae B, the human
CC   MAGE-1 tumour-associated antigen and a hexahistidine tail.  The
CC   invention relates to MAGE proteins fused to an immunological fusion
CC   partner such as LPD.  The LPD moiety provides the fusion protein
CC   with additional exogenous T-cell epitopes and also increase
CC   expression levels in E. coli.  The lipid tail ensures optimal
CC   presentation of the antigen to antigen-presenting cells.  The
CC   affinity tag facilitates purification.  The novel fusion proteins
CC   provide vaccines for immunotherapy of melanomas or other
CC   MAGE-associated tumours like breast, bladder, lung and non-small
CC   cell lung cancer, head and squamous cell carcinoma, colon carcinoma
CC   and oesophagus carcinoma.
XX
SQ   Sequence 1341 BP; 336 A; 327 C; 351 G; 327 T; 0 other;


  Query Match              0.6%;  Score 17;  DB 20;  Length 1341;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1652 ggtcctgggcaccctgg 1668
           |||||||||||||||||
Db     498 ggtcctgggcaccctgg 514


RESULT  50
Q14450/c
ID   Q14450 standard; RNA; 1345 BP.
XX
AC   Q14450;
XX
DT   11-DEC-1991  (first entry)
XX
DE   16S RNA from ATCC 19858 sulphate-reducing bacteria.
XX
KW   SRB; Desulphovibro; Desulphotomaculum; ribosomal RNA; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
XX
PN   US5049489-A.
XX
PD   17-SEP-1991.
XX
PF   17-APR-1989;   89US-0339277.
XX
PR   17-APR-1989;   89US-0339277.
XX
PA   (STAH ) STANDARD OIL CO.
PA   (OHIO ) OHIO OIL CO.
XX
PI   Aldrich KJ,  Brink DE;
XX
DR   WPI; 1991-294983/40.
XX
PT   Assay for sulphate-reducing bacteria - by hybridisation using a
PT   labelled oligo-nucleotide probe corresponding to 16S rRNA of the
PT   bacteria
XX
PS   Disclosure; Fig 1; 16pp; English.
XX
CC   The sequence is shown folded into secondary structure in the
CC   specification.  It was compared with other sequences available in
CC   the literature to design probes specific for SRB.  The probes can
CC   be used for the rapid identification and quantification of SRB in a
CC   sample, e.g. oil-field prodn. waters, water from water treatment
CC   facilities, or samples from the gut of ruminant animals.
CC   See also Q13729-Q13733.
XX
SQ   Sequence 1345 BP; 331 A; 297 C; 435 G; 249 U; 33 other;


  Query Match              0.6%;  Score 17;  DB 12;  Length 1345;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     20 cggctttctcagttttg 36
          |||||||||||||||||
Db   1274 CGGCTTTCTCAGTTTTG 1258


RESULT  51
T27644/c
ID   T27644 standard; cDNA; 1557 BP.
XX
AC   T27644;
XX
DT   14-NOV-1996  (first entry)
XX
DE   cDNA encoding protein for releasing G1 arrest in an animal cell.
XX
KW   G1 arrest; animal cell; human foreskin; cell cycle; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             280..1290
FT                   /*tag= a
FT   Misc-difference 1436
FT                   /*tag= b
FT                   /note= "Given in the specification as 5"
XX
PN   JP08092288-A.
XX
PD   09-APR-1996.
XX
PF   21-SEP-1994;   94JP-0251537.
XX
PR   21-SEP-1994;   94JP-0251537.
XX
PA   (TEIJ ) TEIJIN LTD.
XX
DR   WPI; 1996-236098/24.
DR   P-PSDB; R96248.
XX
PT   DNA encoding G1 arrest-releasing protein - useful for the control of
PT   the cell cycle
XX
PS   Claim 3; Page 9-10; 12pp; Japanese.
XX
CC   This sequence encodes a protein which can release G1 arrest of
CC   an animal cell.  This sequence was isolated from a human foreskin
CC   cDNA library.  The protein can be used in the control of the cell
CC   cycle.
XX
SQ   Sequence 1557 BP; 438 A; 349 C; 363 G; 406 T; 1 other;


  Query Match              0.6%;  Score 17;  DB 17;  Length 1557;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    708 aatgagccacaccttcc 724
          |||||||||||||||||
Db   1014 AATGAGCCACACCTTCC 998


RESULT  52
T98595/c
ID   T98595 standard; DNA; 1591 BP.
```

T98595;

06-NOV-1998 (first entry)

DNA encoding GTP-binding proteins ERA homolog.

Streptococcus pneumoniae protein; genetic immunisation; antagonist; immunological response; inoculation; antibody production; inhibitor; T cell immune response; antimicrobial compound; bacterial adhesion; extracellular matrix protein; protein-mediated cell invasion; wound; pathogenesis; ss.

Streptococcus pneumoniae.

Key	Location/Qualifiers
CDS	complement (592..1086)
	/*tag= a
CDS	complement (1022..1492)
	/*tag= b

WO9743303-A1.

20-NOV-1997.

14-MAY-1997; 97WO-US07950.

14-MAY-1996; 96US-0017670.

(SMIK) SMITHKLINE BEECHAM CORP.
(SMIK) SMITHKLINE BEECHAM PLC.

Black MT, Hodgson JE, Knowles DJC, Nicholas RO;
Stodola RK;

WPI; 1998-008793/01.
P-PSDB; W38537, W38538.

Novel Streptococcus pneumoniae proteins and related DNA - useful for diagnosing anti-microbial agents for treatment of bacterial infections

Claim 4; Pages 129-130; 483pp; English.

This sequence encodes two Streptococcus pneumoniae proteins (based on homology with Streptociccus mutans proteins) are GTP-binding proteins ERA homolog, and represents a DNA sequence of the invention. The DNA sequences were isolated from Streptococcus pneumoniae strain 0100993 (NCIMB 40794). The Streptococcus pneumoniae proteins of the invention can be used to identify compounds which interact with and inhibit or activate the activity of the proteins. Antagonists can be used to treat diseases caused by S. pneumoniae proteins, through genetic immunisation. They can also be used to induce an immunological response in a mammal by inoculation with the S. pneumoniae proteins or delivery of the encoding nucleic acids in a vector adequate to produce antibody and/or T cell immune responses to protect the animal from disease. The proteins can also be used to identify antimicrobial compounds which are capable of inhibiting their bioactivity. In particular the proteins of the invention can be used to prevent adhesion of bacteria to mammalian extracellular matrix proteins on in-dwelling devices or in wounds, to block protein-mediated mammalian cell invasion, and to block the normal progression of pathogenesis in infections initiated other than by the implantation of in-dwelling devices or other surgical techniques.

Sequence 1591 BP; 452 A; 378 C; 313 G; 448 T; 0 other;

```
Query Match             0.6%;  Score 17;  DB 19;  Length 1591;
Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1947 gatttggaagagtttca 1963
           |||||||||||||||||
Db    435 GATTTGGAAGAGTTTCA 419
```

RESULT 53
V42963
ID V42963 standard; DNA; 1592 BP.
XX
AC V42963;
XX
DT 09-NOV-1998 (first entry)
XX
DE Streptococcus pneumoniae polypeptide coding region.
XX
KW Polypeptide; ORF; open reading frame; infection; bacterial;
KW streptococcal; bacteremia; diagnosis; prophylaxis; ds.
XX
OS Streptococcus pneumoniae.
XX

FH	Key	Location/Qualifiers
FT	CDS	573..1001
FT		/*tag= a
FT		/note= "polypeptide"

XX
PN WO9823631-A1.
XX
PD 04-JUN-1998.
XX
PF 24-NOV-1997; 97WO-US21976.
XX
PR 27-NOV-1996; 96US-0031879.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;
PI Reid RH, Zarfos PN;
XX
DR WPI; 1998-322654/28.
DR P-PSDB; W62683.
XX
PT Streptococcus pneumoniae polynucleotides - useful for developing
PT products for diagnosis, prevention and treatment of infections e.g.
PT pneumonia, bacteremia, meningitis or endocarditis
XX
PS Claim 1; Page 56-57; 181pp; English.
XX
CC The sequence is that of a Streptococcal polypeptide coding region.
CC The polypeptide can potentially be used for the diagnosis and
CC prevention of bacterial infections, especially SP infection.
CC It may be used for the treatment of diseases such as otitis media,
CC conjunctivitis, pneumonia, bacteremia, meningitis, sinusitis, pleural
CC empyema, endocarditis or infection of the cerebrospinal fluid.
XX
SQ Sequence 1592 BP; 454 A; 317 C; 376 G; 445 T; 0 other;

```
Query Match             0.6%;  Score 17;  DB 19;  Length 1592;
Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1947 gatttggaagagtttca 1963
           |||||||||||||||||
Db   1158 gatttggaagagtttca 1174
```

RESULT 54
V68059
ID V68059 standard; DNA; 1632 BP.
XX
AC V68059;
XX
DT 02-FEB-1999 (first entry)
XX

DE Neurodegenerative polypeptide HHPDZ65var coding sequence.
XX
KW Neurodegenerative polypeptide; HHPDZ65; stroke; pain; epilepsy; therapy;
KW neurodegenerative disease; ss.
XX
OS Homo sapiens.
XX
PN EP875570-A2.
XX
PD 04-NOV-1998.
XX
PF 15-APR-1998; 98EP-0302912.
XX
PR 19-FEB-1998; 98GB-0003566.
PR 01-MAY-1997; 97GB-0008936.
PR 18-DEC-1997; 97EP-0310289.
XX
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Bingham S, Davis J, Doe TR, Harrison DC, Topp S;
XX
DR WPI; 1998-559436/48.
DR P-PSDB; W80318.
XX
PT HHPDZ65 polypeptide(s), their corresponding DNA, antibodies,
PT agonists and antagonists - are useful in the treatment of stroke,
PT pain, epilepsy and neurodegenerative diseases
XX
PS Claim 21; Page 16; 31pp; English.
XX
CC This sequence encodes the HHPDZ65var neurodegenerative
CC polypeptide of the invention. HHPDZ65 is useful for the treatment of
CC stroke, pain, epilepsy, neurodegenerative diseases and others. The DNAs
CC and proteins are useful in a method for screening to identify compounds
CC which stimulate or inhibit the function of the HHPDZ65 proteins. The
CC polypeptides are useful in a process for diagnosing a disease or a
CC susceptibility to a disease in a subject related to expression or
CC activity of the HHPDZ65 polypeptides.
XX
SQ Sequence 1632 BP; 312 A; 529 C; 497 G; 294 T; 0 other;

Query Match 0.6%; Score 17; DB 19; Length 1632;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2624 cacggtcccccaggagg 2640
||||||||||||||||
Db 1579 cacggtcccccaggagg 1595

RESULT 55
Z25023
ID Z25023 standard; cDNA; 1664 BP.
XX
AC Z25023;
XX
DT 06-DEC-1999 (first entry)
XX
DE Murine D6 encoding cDNA SEQ ID NO:4.
XX
KW D6; G protein-coupled heptahelical receptor; diagnosis; asthma;
KW respiratory inflammatory disorder; identification; ss.
XX
OS Mus sp.
XX
PN WO9947697-A1.
XX
PD 23-SEP-1999.
XX
PF 19-MAR-1999; 99WO-US06075.
XX
PR 20-MAR-1998; 98US-0045583.
XX
PA (MILL-) MILLENNIUM PHARM INC.
PA (CRCT-) CRC TECHNOLOGY LTD.
XX
PI Graham GJ, Benjamin Nibbs RJ, Gonzalo J, Gutierrez-Ramos J;
XX
DR WPI; 1999-562123/47.
DR P-PSDB; Y41682.
XX
PT Identification of D6 G-protein coupled receptor binding compounds and
PT modulators, useful in treatment of asthma -
XX
PS Claim 2; Fig 2; 152pp; English.
XX
CC Methods have been developed for identifying a compound, which binds to a
CC human or murine D6 protein, an allelic variant or a fragment comprises
CC detecting binding of the test compound to the protein. Also described in
CC the present invention are: (1) a method for identifying a compound
CC capable of treating a disorder characterised by aberrant D6 nucleic acid
CC expression of D6 protein activity; (2) a method for treating a subject
CC having a disorder characterized by aberrant D6 protein activity or
CC nucleic acid expression comprising administering to the subject a D6
CC modulator such that treatment of the subject occurs; and (3) methods for
CC identifying a compound that modulates the activity of a Human or murine
CC D6 protein, an allelic variant or a fragment. The methods are useful for
CC identifying compounds capable of treating disorders, especially a
CC respiratory inflammatory disorder, characterized by aberrant D6 nucleic
CC acid expression or D6 protein activity. In particular, the disorder is
CC asthma. D6 modulators are used to treat asthma. The present sequence
CC encodes the murine D6 protein.
XX
SQ Sequence 1664 BP; 313 A; 484 C; 436 G; 431 T; 0 other;

Query Match 0.6%; Score 17; DB 20; Length 1664;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2212 tgctgaaccacttcagc 2228
||||||||||||||||
Db 1479 tgctgaaccacttcagc 1495

RESULT 56
V72116
ID V72116 standard; cDNA; 1668 BP.
XX
AC V72116;
XX
DT 10-MAY-1999 (first entry)
XX
DE Mouse FAST-1 coding region.
XX
KW Transforming growth factor-beta superfamily signalling; modulator; Smad2;
KW TGF-beta; detection; FAST-1; MH2 domain; Smad interaction domain; SID;
KW treatment; developmental; disorder; immunological; cancer; diagnosis; ss.
XX
OS Mus sp.
XX
PN WO9853830-A1.
XX
PD 03-DEC-1998.
XX
PF 28-MAY-1998; 98WO-US10983.
XX
PR 28-MAY-1997; 97US-0047991.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Chen X, Whitman M;
XX
DR WPI; 1999-059773/05.
DR P-PSDB; W90249.

XX
DT Modulating TGF-beta superfamily signalling - comprises use of compounds identified in assays with Smad2, FAST-1 and Smad3, used to develop products for treating, e.g. developmental disorders

XX
RL Example XII; Page 70; 107pp; English.

XX
CC This sequence encodes a mouse FAST-1 protein which is used in a method to detect a compound capable of modulating transforming growth factor-beta (TGF-beta) superfamily signalling. The invention describes a complex which forms between FAST-1 and Smad2 and this complex is specifically induced by signals generated by a TGF-beta superfamily member. A domain of FAST-1 directly interacts with Smad2 and this interaction is mediated by specific domains of the two interacting molecules, namely, the MH2 domain of Smad2 and the Smad interaction domain (SID) of FAST-1. The methods and compounds described are useful for the detection and treatment of conditions involving abnormal TGF-beta superfamily signalling. They can be used to treat e.g. developmental disorders, immunological disorders and cancer. The products can also be used for detection and diagnosis.

XX
SQ Sequence 1668 BP; 368 A; 517 C; 432 G; 351 T; 0 other;

Query Match 0.6%; Score 17; DB 20; Length 1668;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1074 cacatggccccagcatc 1090
||||||||||||||||
Db 621 cacatggccccagcatc 637

RESULT 57
V69719
ID V69719 standard; cDNA; 1691 BP.
XX
AC V69719;
XX
DT 01-MAR-1999 (first entry)
XX
DE Tumour rejection antigen precursor MAGE-A1 cDNA.
XX
KW MAGE-A1; human; tumour rejection antigen precursor; TRAP; therapy; diagnosis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 204..1133
FT /*tag= a
XX
PN WO9849184-A1.
XX
PD 05-NOV-1998.
XX
PF 24-APR-1998; 98WO-US08493.
XX
PR 25-APR-1997; 97US-0845528.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Boon-Falleur T, De Smet C, Lucas S;
XX
DR WPI; 1999-024041/02.
DR P-PSDB; W81548.
XX
PT Tumour rejection antigen precursors - used for determining presence of cytolytic T cells specific for complexes of a human leukocyte antigen
XX
PS Disclosure; Page 46-47; 84pp; English.
XX
CC This nucleotide sequence comprises human tumour rejection antigen precursor (TRAP) MAGE-A1 cDNA, which encodes a 309-amino acid polypeptide (see W81548). MAGE-A1 cDNA shows homology to novel human MAGE-C1 cDNA (see V69720), especially in exons 2 and 3. The open reading frame of MACE-C1, however, is about 2 kb longer than that of MAGE-A1, most of the difference being accounted for by a large repetitive sequence. MAGE-C1 (see W81546) is a novel member of the MAGE family that may be recognised by cytotoxic T cells, leading to lysis of the tumour cells which express it. It is expressed in a variety of tumours and in normal testis cells, but not by other normal cells. The invention provides MACE-C1 and MACE-C2 nucleic acids and polypeptides, useful e.g. in a claimed method for determining the presence of cytolytic T cells specific for complexes of a human leukocyte antigen (HLA).
XX
SQ Sequence 1691 BP; 410 A; 389 C; 465 G; 427 T; 0 other;

Query Match 0.6%; Score 17; DB 20; Length 1691;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1652 ggtcctgggcaccctgg 1668
|||||||||||||||||
Db 317 ggtcctgggcaccctgg 333

RESULT 58
Z52876
ID Z52876 standard; cDNA; 1702 BP.
XX
AC Z52876;
XX
DT 14-MAR-2000 (first entry)
XX
DE Human prostate tumor cDNA library derived EST fragment #19.
XX
KW Pancreas; tumor; EST; expressed sequence tag; human; cytostatic; treatment; ds.
XX
OS Homo sapiens.
XX
PN DE19820190-A1.
XX
PD 04-NOV-1999.
XX
PF 28-APR-1998; 98DE-1020190.
XX
PR 28-APR-1998; 98DE-1020190.
XX
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX
DR WPI; 1999-621386/54.
DR P-PSDB; Y73868, Y73869, Y73870.
XX
PT New human nucleic acid sequences from pancreatic tumors, and related proteins -
XX
PS Claim 2; Page 197-198; 502pp; German.
XX
CC This invention describes novel polypeptides and their encoding nucleic acids derived from human pancreatic tumor tissue which have cytostatic activity. The sequences are also useful in producing pharmaceutical compositions for treatment of pancreatic tumors. Z52858-Z53014 represent expressed sequence tag (EST) fragments derived from a human pancreatic tumor cDNA library and which encode the proteins represented in Y73814-Y74252.
XX
SQ Sequence 1702 BP; 371 A; 517 C; 476 G; 338 T; 0 other;

```
  Query Match             0.6%;  Score 17;  DB 20;  Length 1702;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1341 agggagtggcagaggga 1357
          |||||||||||||||||
Db   1112 agggagtggcagaggga 1128


RESULT  59
V38385
ID   V38385 standard; cDNA; 1724 BP.
XX
AC   V38385;
XX
DT   24-NOV-1998  (first entry)
XX
DE   Beta(1 -> 4)-N-acetylglucosaminyl-transferase (GnT-IV)b encoding cDNA.
XX
KW   Beta(1 -> 4)-N-acetylglucosaminyl-transferase; GnT-IV; bovine; human;
KW   enzyme; sugar chain subunit; branched oligosaccharide; polysaccharide;
KW   drug; reagent; food; biopolymer; glycoprotein; erythropoietin; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             43..1689
FT                   /*tag= a
FT                   /product= "GnT-IVb enzyme"
XX
PN   WO9826053-A1.
XX
PD   18-JUN-1998.
XX
PF   10-DEC-1997;   97WO-JP04546.
XX
PR   18-JUN-1997;   97JP-0161462.
PR   12-DEC-1996;   96JP-0332411.
XX
PA   (KIRI ) KIRIN BEER KK.
XX
PI   Minowa M,  Oguri S,  Takeuchi M,  Taniguchi N,  Yoshida A;
XX
DR   WPI; 1998-348516/30.
DR   P-PSDB; W63559.
XX
PT   Recombinant beta(1-4)N-acetyl:glucosaminyl:transferase - allows
PT   production of difficultly accessible branched poly:saccharides for
PT   food and drug use
XX
PS   Claim 16; Pages 70-74; 112pp; Japanese.
XX
CC   This cDNA encodes a human beta(1 -> 4)-N-acetylglucosaminyl-transferase
CC   (GnT-IV)b enzyme. The invention provides bovine and human GnT-IV enzymes
CC   that can be used for converting sugar chain subunits having one structure
CC   to another structure. Vectors containing the DNA sequences encoding these
CC   enzymes can be used to transform host cells for the production of the
CC   GnT-IV enzymes. The enzymes are useful in the production of branched
CC   oligosaccharides and polysaccharides which are difficult of access by
CC   other methods. They are also useful in the production of drugs, reagents
CC   and foods and in modifying the properties of biopolymers containing sugar
CC   chains. The enzyme may also be used for the preparation of glycoproteins
CC   such as erythropoietin.
XX
SQ   Sequence 1724 BP; 353 A; 553 C; 501 G; 317 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 19;  Length 1724;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1341 agggagtggcagaggga 1357
          |||||||||||||||||
Db   1565 agggagtggcagaggga 1581


RESULT  60
X85940
ID   X85940 standard; DNA; 1816 BP.
XX
AC   X85940;
XX
DT   13-SEP-1999  (first entry)
XX
DE   DNA encoding human cell division regulator (HCDR) 1.
XX
KW   Human cell division regulator; HCDR; interphase; inflammation;
KW   cell proliferation; apoptosis; neurodeficiency;
KW   neurodegenerative disease; aplastic anaemia; ischaemic injury;
KW   liver damage; viral infection; hepatitis B; hepatitis C; ss.
XX
OS   Homo sapiens.
XX
PN   US5928899-A.
XX
PD   27-JUL-1999.
XX
PF   01-OCT-1998;   98US-0165234.
XX
PR   15-OCT-1997;   97US-0951148.
PR   01-OCT-1998;   98US-0165234.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Bandman O,  Corley NC,  Hillman JL,  Lal P,  Shah P;
XX
DR   WPI; 1999-429499/36.
DR   P-PSDB; Y23782.
XX
PT   Cell division regulators active in interphase
XX
PS   Example 1; Fig 1A-E; 59pp; English.
XX
CC   The present sequence encodes human cell division regulator (HCDR) 1.
CC   HCDR proteins are active in interphase, and are used for the
CC   treatment or prevention of inflammation and disorders associated with
CC   cell proliferation and apoptosis. HCDR may be administered to a
CC   patient having a disorder associated with an increase in apoptosis,
CC   such a disorder may be e.g. neurodeficiency, a neurodegenerative
CC   disease, aplastic anaemia, an ischaemic injury, liver damage, or a
CC   viral infection such as hepatitis B or C.
XX
SQ   Sequence 1816 BP; 494 A; 473 C; 494 G; 355 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 20;  Length 1816;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    804 ggaaacttcttggtgct 820
          |||||||||||||||||
Db    717 ggaaacttcttggtgct 733


RESULT  61
X01577
ID   X01577 standard; DNA; 1818 BP.
XX
AC   X01577;
XX
DT   04-MAY-1999  (first entry)
XX
DE   Human HCDR-1 coding sequence.
XX
KW   Human; HCDR-1; HCDR-2; HCDR-3; human cell division regulator; apoptosis;
KW   inflammation; cell proliferation disorder; adenocarcinoma; AIDS; ss.
```

```
XX
OS   Homo sapiens.
XX
PN   US5871973-A.
XX
PD   16-FEB-1999.
XX
PF   15-OCT-1997;   97US-0951148.
XX
PR   15-OCT-1997;   97US-0951148.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Bandman O,  Corley NC,  Hillman JL,  Lal P,  Shah P;
XX
DR   WPI; 1999-166646/14.
DR   P-PSDB; W73971.
XX
PT   New polynucleotides encoding human cell division regulators (HCDR)
PT   - useful for diagnosing, preventing and treating inflammation and
PT   disorders associated with cell proliferation and apoptosis
XX
PS   Claim 4; Fig 1; 59pp; English.
XX
CC   This sequence encodes the human cell division regulator-1 (HCDR-1)
CC   protein of the invention. Polynucleotides complementary to the HCDR-1
CC   coding sequence can be used as probes to detect the DNA in a sample. The
CC   polynucleotide sequences encoding HCDR may be used to prevent/treat
CC   inflammation and disorders associated with cell proliferation and
CC   apoptosis and in assays that detect activation of cancers.
CC   Polynucleotides encoding HCDR may be used for the diagnosis of conditions
CC   associated with expression of HCDR, including disorders associated with
CC   cell proliferation/apoptosis e.g. adenocarcinoma and AIDS. The
CC   polynucleotides may also be used in Southern or Northern analysis, dot
CC   blot, or other membrane based technologies; in PCR technologies; or in
CC   dipstick, pin, or ELISA assays or microarrays utilising fluids or tissues
CC   from patient biopsies to detect altered HCDR expression.
XX
SQ   Sequence 1818 BP; 494 A; 473 C; 494 G; 355 T; 2 other;


  Query Match             0.6%;  Score 17;  DB 20;  Length 1818;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    804 ggaaacttcttggtgct 820
          |||||||||||||||||
Db    719 ggaaacttcttggtgct 735


RESULT  62
A15550/c
ID   A15550 standard; cDNA to mRNA; 1906 BP.
XX
AC   A15550;
XX
DT   31-JUL-2000  (first entry)
XX
DE   Human TRAF four associated factor TFAF2 coding sequence.
XX
KW   TFAF2; TRAF four associated factor 2; tumour formation; breast cancer;
KW   TRAF4; TNF receptor associated factor; tumour diagnosis; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             13..1230
FT                   /*tag=  a
FT                   /product= TFAF2
XX
PN   CA2245340-A1.
XX
PD   19-FEB-2000.
XX
PF   19-AUG-1998;   98CA-2245340.
XX
PR   19-AUG-1998;   98CA-2245340.
XX
PA   (MEDI-) MEDICAL & BIOLOGICAL LAB CO LTD.
XX
PI   Yano M,  Toji S,  Tamai K;
XX
DR   WPI; 2000-351124/31.
DR   P-PSDB; Y94209.
XX
PT   Novel tumour necrosis factor receptor associated factor 4 associated
PT   factors useful for developing cancer screens, and treating tumours -
XX
PS   Claim 4; Page 47-50; 68pp; English.
XX
CC   The present sequence is the coding sequence of human TRAF four
CC   associated factor TFAF2. The gene was discovered by screening a human
CC   placenta cDNA library using a two-hybrid system. The protein associates
CC   with the TRAF domain located at the carboxyl-terminal of TNF
CC   receptor associated factor 4 (TRAF4), which is believed to be an
CC   oncoprotein. Antibodies that bind to TNF four associated factors (TFAFs)
CC   may be used to treat or diagnose tumours (e.g. breast cancer) when
CC   labelled with an isotope or an appropriate drug, precursor or enzyme.
CC   Antagonists, agonists, and antisense sequences of TFAFs may be used to
CC   treat cancers. TFAF proteins, antibodies that recognise them and DNAs
CC   for them may be useful as tools for cancer research.
XX
SQ   Sequence 1906 BP; 655 A; 320 C; 372 G; 559 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 21;  Length 1906;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2796 gtttttaaagaagtctt 2812
           |||||||||||||||||
Db     582 GTTTTTAAAGAAGTCTT 566


RESULT  63
T09084/c
ID   T09084 standard; cDNA to mRNA; 2160 BP.
XX
AC   T09084;
XX
DT   09-MAY-1996  (first entry)
XX
DE   Nitrite reductase gene.
XX
KW   nitrite reductase; transgenic crop; transgenic tree; detoxification;
KW   atmospheric pollutant; ds.
XX
OS   Populus nigra L. var italica.
XX
FH   Key             Location/Qualifiers
FT   CDS             40..1806
FT                   /*tag=  a
FT                   /product= nitrite_reductase
FT   primer_bind     367..390
FT                   /*tag=  b
FT   primer_bind     838..860
FT                   /*tag=  c
FT   polyA_signal    2114..2119
FT                   /*tag=  d
FT   misc_feature    2146..2160
FT                   /*tag=  e
FT                   /label= poly_A_site
XX
PN   JP07236486-A.
XX
PD   12-SEP-1995.
```

```
AX
PF   02-MAR-1994;   94JP-0032359.
AX
PR   02-MAR-1994;   94JP-0032359.
XX
PA   (TOYT ) TOYOTA JIDOSHA KK.
XX
DR   WPI; 1995-347454/45.
DR   P-PSDB; R87973.
XX
PT   Nitrite reductase gene from Populus nigra - useful for breeding
PT   trees to remove nitrogen di:oxide from the atmosphere or for crops
PT   which produce fewer carcinogenic nitrosamine(s)
XX
PS   Claim 1; Page 5-8; 8pp; Japanese.
XX
CC   The gene encodes a nitrite reductase gene isolated from Populus nigra.
CC   The gene is useful in generation of street trees with a high power for
CC   clarfiying NO2, an atmospheric pollutant.  The gene can also be used in
CC   transgenic plants, esp. crops, bred to contain less nitrosamines (a
CC   carcinogenic substance present in food).
XX
SQ   Sequence 2160 BP; 660 A; 388 C; 570 G; 542 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 16;  Length 2160;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1711 gcttgccaagtatcttg 1727
           |||||||||||||||||
Db    495 GCTTGCCAAGTATCTTG 479


RESULT  64
Z50582/c
ID   Z50582 standard; cDNA; 2284 BP.
XX
AC   Z50582;
XX
DT   23-MAY-2000  (first entry)
XX
DE   Human epidermal protein-6 cDNA.
XX
KW   Human epidermal protein-6; HEPI; epithelial disorder; scabies;
KW   dyshidrotic eczema; cell proliferative disorder; actinic keratosis;
KW   arteriosclerosis; autoimmune disorder; inflammatory disorder;
KW   acquired immune deficiency syndrome; AIDS; Addison's disease; antiHIV;
KW   dermatological; anitarteriosclerotic; antiinflammatory;
KW   immunosuppressive; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             47..1636
FT                   /*tag=  a
FT                   /product= "Human epidermal protein-6"
FT   sig_peptide     1..142
FT                   /*tag=  b
FT   mat_peptide     148..1636
FT                   /*tag=  c
FT                   /product= "Mature human epidermal protein-6"
XX
PN   WO200006727-A2.
XX
PD   10-FEB-2000.
XX
PF   27-JUL-1999;   99WO-US17107.
XX
PR   28-JUL-1998;   98US-0155203.
PR   07-DEC-1998;   98US-0155254.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Tang YT,  Lal P,  Corley NC,  Guegler KJ,  Patterson C,  Baughn MR;
PI   Yue H;
XX
DR   WPI; 2000-195295/17.
DR   P-PSDB; Y44989.
XX
PT   New human epidermal proteins (HEPI-1) to (HEPI-6) useful for the
PT   diagnosis, treatment and prevention of epithelial, cell proliferative,
PT   and autoimmune inflammatory disorders -
XX
PS   Claim 7; Page 78-79; 82pp; English.
XX
CC   The present cDNA sequence encodes human epidermal protein-6 (HEPI). The
CC   cDNA clone is derived from PTHYNOT03 library which was constructed using
CC   RNA isolated from left parathyroid tissue of a 69-year-old caucasian
CC   female during a partial parathyroidectomy. Recombinant vectors
CC   comprising HEPI cDNA are introduced into host cells for protein
CC   expression. The HEPI proteins are useful for the treatment of epithelial
CC   disorders, including dyshidrotic eczema and scabies, cell proliferative
CC   disorders including actinic keratosis and arteriosclerosis, and
CC   autoimmune/inflammatory disorders like acquired immune deficiency
CC   syndrome (AIDS) and Addison's disease. Pharmaceutical compositions
CC   comprising HEPI proteins are useful for treating disorders associated
CC   with altered HEPI expression.
XX
SQ   Sequence 2284 BP; 452 A; 775 C; 534 G; 523 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 21;  Length 2284;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1881 caggaaggggctgagat 1897
           |||||||||||||||||
Db   2262 CAGGAAGGGGCTGAGAT 2246


RESULT  65
X84103
ID   X84103 standard; DNA; 2418 BP.
XX
AC   X84103;
XX
DT   08-SEP-1999  (first entry)
XX
DE   E antigen precursor gene.
XX
KW   Tumour rejection antigen; vaccine; cancer; E antigen precursor gene; ss.
XX
OS   Homo sapiens.
XX
PN   US5925729-A.
XX
PD   20-JUL-1999.
XX
PF   02-MAY-1994;   94US-0142368.
XX
PR   02-MAY-1994;   94US-0142368.
PR   23-MAY-1991;   91US-0705702.
PR   09-JUL-1991;   91US-0728838.
PR   23-SEP-1991;   91US-0764365.
PR   12-DEC-1991;   91US-0807043.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
XX
PI   Boon T,  Chomez P,  De Plaen E,  Lurquin C,  Traversari C;
PI   Van Den Eynde B, Van Der Bruggen P,  Van Pel A;
XX
DR   WPI; 1999-418294/35.
XX
PT   New tumour rejection antigen is useful as a vaccine against
PT   cancerous diseases
```

XX
RS Example 20; Column 15-18; 58pp; English.
XX
CC This sequence represents the E antigen precursor gene.
CC The invention relates to a tumour rejection antigen sequence that is
CC useful as a tumour rejection antigen for vaccination against cancerous
CC conditions.
XX
SQ Sequence 2418 BP; 562 A; 582 C; 675 G; 599 T; 0 other;

Query Match 0.6%; Score 17; DB 20; Length 2418;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1652 ggtcctgggcaccctgg 1668
||||||||||||||||
Db 738 ggtcctgggcaccctgg 754

RESULT 66
Q32351
ID Q32351 standard; DNA; 2419 BP.
XX
AC Q32351;
XX
DT 22-APR-1993 (first entry)
XX
DE Antigen E gene.
XX
KW Stable; antigen; E; D; F; A; human; melanoma; cell line; MZ2-MEL;
KW cytolytic T cell; MEL3.1; open reading frame; homology; MAGE;
KW melanoma antigen; ss.
XX
OS Homo sapiens.
XX
PN WO9220356-A.
XX
PD 26-NOV-1992.
XX
PF 22-MAY-1992; 92WO-US04354.
XX
PR 23-MAY-1991; 91US-0705702.
PR 09-JUL-1991; 91US-0728838.
PR 23-SEP-1991; 91US-0764364.
PR 12-DEC-1991; 91US-0807043.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Boon T, Chomez P, De Plaen E, Lurquin C, Traversari C;
PI Van Den Eynde B, Van Der Bruggen P, Van Pel A;
XX
DR WPI; 1992-415460/50.
XX
PT Nucleic acid mol. encoding a human tumour rejection antigen
PT precursor - useful as an immunostimulant in a vaccine for
PT treating and preventing cancers, also useful in diagnosis
XX
PS Disclosure; Page 69-70; 142pp; English.
XX
CC This sequence encodes the stable antigen E. This antigen is expressed
CC along with antigens "D, F and A" by the human melanoma cell line MZ2-
CC MEL. These antigens are all recognised by cytolytic T cells. A
CC subline of MZ2-MEL is MEL3.1 which only expresses antigen E. This
CC cell line was chosen as a source for the isolation of this sequence.
CC This sequence was found to contain three exons. The open reading frame
CC for antigen E was contained within the first two exons. An ATG is
CC located at position 66 of exon 3 and is followed by an 828 base pair
CC reading frame. The three exons contain 65, 73 and 1551 base pairs.
CC During the isolation of this sequence two different but closely reated
CC cDNAs were also identified. These cDNAs, when tested, did not
CC transfer expression of antigen E, but they did show substantial
CC homology to the antigen E cDNA sequence. These new cDNAs represent a
CC new family of genes refered to as melanoma antigens (MAGE) (see also
CC Q32352-69).
XX
SQ Sequence 2419 BP; 562 A; 581 C; 677 G; 599 T; 0 other;

Query Match 0.6%; Score 17; DB 13; Length 2419;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1652 ggtcctgggcaccctgg 1668
||||||||||||||||
Db 739 ggtcctgggcaccctgg 755

RESULT 67
Q72476
ID Q72476 standard; DNA; 2419 BP.
XX
AC Q72476;
XX
DT 21-JUN-1995 (first entry)
XX
DE Tumour rejection antigen E encoding DNA.
XX
KW Tumour rejection antigen E; melanoma antigen-3; MAGE-3;
KW cancer; cytolytic T cells; antigen D; human leucocyte antigen; ss.
XX
OS Homo sapiens.
XX
PN WO9423031-A.
XX
PD 13-OCT-1994.
XX
PF 17-MAR-1994; 94WO-US02877.
XX
PR 26-MAR-1993; 93US-0037230.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Boon-falleur T, Gaugler B, Van DEN EYNDE B, Van DER BRUGGEN P;
XX
DR WPI; 1994-333192/41.
XX
PT New tumour rejection antigen precursor MAGE3 - useful in
PT treatment and diagnosis of cancer
XX
PS Disclosure; Page 58; 105pp; English.
XX
CC Q72476 encodes tumour rejection antigen E, another sequence
CC Q72470 encodes melanoma antigen-3 (MAGE-3) a tumour rejection
CC antigen precursor. Melanomas characterised by the expression of
CC MAGE-3 can be detected, or monitored, by contacting a test sample
CC with an agent that can recognise MAGE-3. The melanoma can be treated
CC by the administration of cytolytic T cells specific for the complex of
CC antigen D (the mature rejection antigen derived from MAGE-3) and a human
CC leucocyte antigen (esp. HLA-A1).
XX
SQ Sequence 2419 BP; 562 A; 581 C; 677 G; 599 T; 0 other;

Query Match 0.6%; Score 17; DB 15; Length 2419;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1652 ggtcctgggcaccctgg 1668
||||||||||||||||
Db 739 ggtcctgggcaccctgg 755

RESULT 68
T05086
ID T05086 standard; DNA; 2419 BP.

XX
AC T05086;
XX
DT 26-FEB-1996 (first entry)
XX
DE MZ2-MEL antigen E precursor gene.
XX
KW Melanoma; MZ2-MEL; tumour rejection antigen; cancer; diagnosis; ss.
XX
OS Homo sapiens.
XX
PN WO9523874-A1.
XX
PD 08-SEP-1995.
XX
PF 23-FEB-1995; 95WO-US02203.
XX
PR 30-NOV-1994; 94US-0346774.
PR 01-MAR-1994; 94US-0204727.
PR 10-MAR-1994; 94US-0209172.
PR 01-SEP-1994; 94US-0299849.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Boon-Falleur T, Brasseur F, Chomez P, De Plaen E;
PI De Smet C, Gaugler B, Lethe B, Marchand M, Patard J;
PI Szikora J, Van Den Eynde B, Van Derbruggen P, Weynants P;
XX
DR WPI; 1995-320586/41.
XX
PT Determn. of cancerous condition(s) - using a nucleic acid as a
PT primer to determine expression of a MAGE tumour rejection antigen
PT precursor
XX
PS Example 20; Page 69-70; 121pp; English.
XX
CC A gene sequence (T05086) hybridizes with a 2.4 kb fragment from
CC human melanoma cell line MZ2-MEL but not with E- antigen loss
CC variants of MZ2-MEL. This E precursor antigen gene sequence was
CC obtd. from a cosmid derived from DNA of the E+ subclone MZ2-MEL 43.
XX
SQ Sequence 2419 BP; 560 A; 581 C; 677 G; 601 T; 0 other;

Query Match 0.6%; Score 17; DB 16; Length 2419;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1652 ggtcctgggcaccctgg 1668
||||||||||||||||
Db 739 ggtcctgggcaccctgg 755

RESULT 69
X84112
ID X84112 standard; DNA; 2419 BP.
XX
AC X84112;
XX
DT 08-SEP-1999 (first entry)
XX
DE Antigen E coding sequence.
XX
KW Tumour rejection antigen; vaccine; cancer; antigen E; ss.
XX
OS Homo sapiens.
XX
PN US5925729-A.
XX
PD 20-JUL-1999.
XX
PF 02-MAY-1994; 94US-0142368.
XX
PR 02-MAY-1994; 94US-0142368.
PR 23-MAY-1991; 91US-0705702.
PR 09-JUL-1991; 91US-0728838.
PR 23-SEP-1991; 91US-0764365.
PR 12-DEC-1991; 91US-0807043.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Boon T, Chomez P, De Plaen E, Lurquin C, Traversari C;
PI Van Den Eynde B, Van Der Bruggen P, Van Pel A;
XX
DR WPI; 1999-418294/35.
XX
PT New tumour rejection antigen is useful as a vaccine against
PT cancerous diseases
XX
PS Disclosure; Column 37-40; 58pp; English.
XX
CC This sequence represents the antigen E coding sequence.
CC The invention relates to a tumour rejection antigen sequence that is
CC useful as a tumour rejection antigen for vaccination against cancerous
CC conditions.
XX
SQ Sequence 2419 BP; 562 A; 581 C; 677 G; 599 T; 0 other;

Query Match 0.6%; Score 17; DB 20; Length 2419;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1652 ggtcctgggcaccctgg 1668
||||||||||||||||
Db 739 ggtcctgggcaccctgg 755

RESULT 70
Q72472
ID Q72472 standard; DNA; 2420 BP.
XX
AC Q72472;
XX
DT 21-JUN-1995 (first entry)
XX
DE Tumour rejection antigen E precursor gene DNA.
XX
KW Tumour antigen rejection precursor E; melanoma antigen-3; MAGE-3;
KW cancer; cytolytic T cells; antigen D; human leucocyte antigen; ss.
XX
OS Homo sapiens.
XX
PN WO9423031-A.
XX
PD 13-OCT-1994.
XX
PF 17-MAR-1994; 94WO-US02877.
XX
PR 26-MAR-1993; 93US-0037230.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Boon-falleur T, Gaugler B, Van DEN EYNDE B, Van DER BRUGGEN P;
XX
DR WPI; 1994-333192/41.
XX
PT New tumour rejection antigen precursor MAGE3 - useful in
PT treatment and diagnosis of cancer
XX
PS Example 20; Page 28; 105pp; English.
XX
CC Q72472 is the tumour rejection antigen E precursor gene, another
CC gene Q72470 encodes melanoma antigen-3 (MAGE-3) also a tumour rejection
CC antigen precursor. Melanomas characterised by the expression of MAGE-3
CC can be detected, or monitored, by contacting a test sample with an

agent that can recognise MAGE-3. The melanoma can be treated by the administration of cytolytic T cells specific for the complex of antigen D (the mature rejection antigen derived from MAGE-3) and a human leucocyte antigen (esp. HLA-A1).

Sequence 2420 BP; 562 A; 582 C; 677 G; 599 T; 0 other;

Query Match 0.6%; Score 17; DB 15; Length 2420;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1652 ggtcctgggcaccctgg 1668
||||||||||||||||||
739 ggtcctgggcaccctgg 755

RESULT 71
Q85435
Q85435 standard; DNA; 2420 BP.

Q85435;

09-OCT-1995 (first entry)

Human melanoma antigen MAGE-1.

Human melanoma antigen; MAGE-1; vaccines; MAGE associated tumours; HLA-restricted cytotoxic T-lymphocyte activity; ss.

Homo sapiens.

Key	Location/Qualifiers
CDS	626..1555
	/*tag= a

WO9504542-A.

16-FEB-1995.

02-AUG-1994; 94WO-US08721.

06-AUG-1993; 93US-0103623.

(CYTE-) CYTEL CORP.

Fikes JD, Livingston BD, Sette AD, Sidney JC;

WPI; 1995-090681/12.
P-PSDB; R70909.

Human melanoma antigen, MAGE-1, peptide(s) - useful for stimulating immune response against melanoma

Example 1; Fig 1; 59pp; English.

Q85435 encodes R70909 human melanoma antigen MAGE-1, it was used to produce the C-terminal MAGE-1 peptides described in R70915 to R70969. These peptides are useful for defining epitopes that engender a HLA-restricted cytotoxic lymphocyte activity against MAGE-1 antigens. Compsns. containing these peptides can be administered, as a vaccine to patients susceptable to MAGE associated tumours, e.g. melanomas.

Sequence 2420 BP; 562 A; 582 C; 677 G; 599 T; 0 other;

Query Match 0.6%; Score 17; DB 16; Length 2420;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1652 ggtcctgggcaccctgg 1668
||||||||||||||||||
Db 739 ggtcctgggcaccctgg 755

RESULT 72
V59595
ID V59595 standard; DNA; 2503 BP.
XX
AC V59595;
XX
DT 06-JAN-1999 (first entry)
XX
DE Human secreted protein gene 85 clone HSDFV29.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.
XX
PN WO9839448-A2.
XX
PD 11-SEP-1998.
XX
PF 06-MAR-1998; 98WO-US04493.
XX
PR 02-OCT-1997; 97US-0061060.
PR 07-MAR-1997; 97US-0038621.
PR 07-MAR-1997; 97US-0040161.
PR 07-MAR-1997; 97US-0040162.
PR 07-MAR-1997; 97US-0040163.
PR 07-MAR-1997; 97US-0040333.
PR 07-MAR-1997; 97US-0040334.
PR 07-MAR-1997; 97US-0040336.
PR 07-MAR-1997; 97US-0040626.
PR 11-APR-1997; 97US-0043311.
PR 11-APR-1997; 97US-0043312.
PR 11-APR-1997; 97US-0043313.
PR 11-APR-1997; 97US-0043314.
PR 11-APR-1997; 97US-0043568.
PR 11-APR-1997; 97US-0043569.
PR 11-APR-1997; 97US-0043576.
PR 11-APR-1997; 97US-0043578.
PR 11-APR-1997; 97US-0043580.
PR 11-APR-1997; 97US-0043669.
PR 11-APR-1997; 97US-0043670.
PR 11-APR-1997; 97US-0043671.
PR 11-APR-1997; 97US-0043672.
PR 11-APR-1997; 97US-0043674.
PR 23-MAY-1997; 97US-0047492.
PR 23-MAY-1997; 97US-0047500.
PR 23-MAY-1997; 97US-0047501.
PR 23-MAY-1997; 97US-0047502.
PR 23-MAY-1997; 97US-0047503.
PR 23-MAY-1997; 97US-0047581.
PR 23-MAY-1997; 97US-0047582.
PR 23-MAY-1997; 97US-0047583.
PR 23-MAY-1997; 97US-0047584.
PR 23-MAY-1997; 97US-0047585.
PR 23-MAY-1997; 97US-0047586.
PR 23-MAY-1997; 97US-0047587.
PR 23-MAY-1997; 97US-0047588.
PR 23-MAY-1997; 97US-0047589.
PR 23-MAY-1997; 97US-0047590.
PR 23-MAY-1997; 97US-0047592.
PR 23-MAY-1997; 97US-0047593.
PR 23-MAY-1997; 97US-0047594.
PR 23-MAY-1997; 97US-0047595.
PR 23-MAY-1997; 97US-0047596.

PR 23-MAY-1997; 97US-0047597.
PR 23-MAY-1997; 97US-0047598.
PR 23-MAY-1997; 97US-0047599.
PR 23-MAY-1997; 97US-0047600.
PR 23-MAY-1997; 97US-0047601.
PR 23-MAY-1997; 97US-0047612.
PR 23-MAY-1997; 97US-0047613.
PR 23-MAY-1997; 97US-0047614.
PR 23-MAY-1997; 97US-0047615.
PR 23-MAY-1997; 97US-0047617.
PR 23-MAY-1997; 97US-0047618.
PR 23-MAY-1997; 97US-0047632.
PR 23-MAY-1997; 97US-0047633.
PR 06-JUN-1997; 97US-0048964.
PR 06-JUN-1997; 97US-0048974.
PR 13-JUN-1997; 97US-0049610.
PR 08-JUL-1997; 97US-0051926.
PR 16-JUL-1997; 97US-0052874.
PR 18-AUG-1997; 97US-0055724.
PR 22-AUG-1997; 97US-0056630.
PR 22-AUG-1997; 97US-0056631.
PR 22-AUG-1997; 97US-0056632.
PR 22-AUG-1997; 97US-0056636.
PR 22-AUG-1997; 97US-0056637.
PR 22-AUG-1997; 97US-0056662.
PR 22-AUG-1997; 97US-0056664.
PR 22-AUG-1997; 97US-0056845.
PR 22-AUG-1997; 97US-0056862.
PR 22-AUG-1997; 97US-0056864.
PR 22-AUG-1997; 97US-0056872.
PR 22-AUG-1997; 97US-0056874.
PR 22-AUG-1997; 97US-0056875.
PR 22-AUG-1997; 97US-0056876.
PR 22-AUG-1997; 97US-0056877.
PR 22-AUG-1997; 97US-0056878.
PR 22-AUG-1997; 97US-0056879.
PR 22-AUG-1997; 97US-0056880.
PR 22-AUG-1997; 97US-0056881.
PR 22-AUG-1997; 97US-0056882.
PR 22-AUG-1997; 97US-0056884.
PR 22-AUG-1997; 97US-0056886.
PR 22-AUG-1997; 97US-0056887.
PR 22-AUG-1997; 97US-0056888.
PR 22-AUG-1997; 97US-0056889.
PR 22-AUG-1997; 97US-0056892.
PR 22-AUG-1997; 97US-0056893.
PR 22-AUG-1997; 97US-0056894.
PR 22-AUG-1997; 97US-0056903.
PR 22-AUG-1997; 97US-0056908.
PR 22-AUG-1997; 97US-0056909.
PR 22-AUG-1997; 97US-0056910.
PR 22-AUG-1997; 97US-0056911.
PR 05-SEP-1997; 97US-0057650.
PR 05-SEP-1997; 97US-0057669.
PR 05-SEP-1997; 97US-0057761.
PR 12-SEP-1997; 97US-0058785.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;
PI Feng P, Ferrie AM, Fischer CL, Florence KA, Greene JM, Hu JS;
PI Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;
PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
XX
DR WPI; 1998-506364/43.
DR P-PSDB; W74815.
XX
PT New isolated human genes and the secreted polypeptide(s) they encode
PT - useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX
PS Claim 1; Page 316-317; 721pp; English.
XX
CC This sequence represents a nucleic acid molecule designated Gene 85 from
CC the human cDNA clone HSDFV29 (deposited as clone ATCC 209076) which
CC encodes a secreted human protein. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. V59502) for increasing the stability of the fused protein
CC as compared to the human protein only.
CC The invention relates to 186 novel genes and their fragments (nucleic
CC acid sequences: V59511-V59812; amino acid sequences W74731-W75026) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 186 polynucleotides, based on
CC which tissues they are most highly expressed in (see V59511 for described
CC uses).
XX
SQ Sequence 2503 BP; 561 A; 705 C; 658 G; 568 T; 11 other;

Query Match 0.6%; Score 17; DB 19; Length 2503;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2444 gctggcaggcggcctgg 2460
||||||||||||||||
Db 583 gctggcaggcggcctgg 599

RESULT 73
Z42096
ID Z42096 standard; cDNA; 2646 BP.
XX
AC Z42096;
XX
DT 31-JAN-2000 (first entry)
XX
DE Human endometrium tumour cDNA derived EST 116.
XX
KW Endometrium; human; tumour; cancer; anticancer; cytostatic; EST:
KW treatment; uterine; gene therapy; expressed sequence tag; ss.
XX
OS Homo sapiens.
XX
PN DE19817948-A1.
XX
PD 21-OCT-1999.
XX
PF 17-APR-1998; 98DE-1017948.
XX
PR 17-APR-1998; 98DE-1017948.
XX
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX
DR WPI; 1999-591957/51.
DR P-PSDB; Y60295, Y60296, Y60297.
XX
PT New nucleic acid sequences expressed in uterine cancer tissues, and
PT derived polypeptides, for treatment of uterine and endometrial cancer
PT and identification of therapeutic agents
XX
PS Claim 3; Page 253; 444pp; German.
XX
CC This invention describes novel human nucleic acid (cDNA) sequences (A),
CC that are highly expressed in uterine tumour tissue and which have
CC anticancer and cytostatic activity. (A) are used (i) for recombinant
CC expression of polypeptides (B) and (ii) to isolate complete genes. (B)
CC are used (i) to identify agents suitable for treatment of uterine or
CC endometrial cancer; (ii) directly for treating these forms of cancer
CC (including expression from gene therapy vectors) and (iii) for generation
CC of specific antibodies. (A) are identified by assembling ESTs (expressed
CC sequence tags) from a particular tissue type before comparison of

CC expression patterns. This allows a significantly longer fragment of the
CC gene to be revealed, so should reduce the number of failures associated
CC with the fact that ESTs from different libraries may represent different
CC parts of the same unknown gene, distorting the estimated frequency of
CC occurrence in a particular tissue. Z41981-Z42121 represent EST fragments
CC derived from a human endometrium tumour cDNA library which encode the
CC protein sequences represented in Y59941-Y60328.
XX
SQ Sequence 2646 BP; 787 A; 502 C; 546 G; 811 T; 0 other;

Query Match 0.6%; Score 17; DB 20; Length 2646;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1463 agccccagcagagaaaa 1479
||||||||||||||||
Db 1717 agccccagcagagaaaa 1733

RESULT 74
V68056
ID V68056 standard; DNA; 2711 BP.
XX
AC V68056;
XX
DT 02-FEB-1999 (first entry)
XX
DE Neurodegenerative polypeptide HHPDZ65 coding sequence.
XX
KW Neurodegenerative polypeptide; HHPDZ65; stroke; pain; epilepsy; therapy;
KW neurodegenerative disease; ss.
XX
OS Homo sapiens.
XX
PN EP875570-A2.
XX
PD 04-NOV-1998.
XX
PF 15-APR-1998; 98EP-0302912.
XX
PR 19-FEB-1998; 98GB-0003566.
PR 01-MAY-1997; 97GB-0008936.
PR 18-DEC-1997; 97EP-0310289.
XX
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Bingham S, Davis J, Doe TR, Harrison DC, Topp S;
XX
DR WPI; 1998-559436/48.
DR P-PSDB; W80315.
XX
PT HHPDZ65 polypeptide(s), their corresponding DNA, antibodies,
PT agonists and antagonists - are useful in the treatment of stroke,
PT pain, epilepsy and neurodegenerative diseases
XX
PS Claim 7; Page 14; 31pp; English.
XX
CC This sequence encodes the HHPDZ65 neurodegenerative polypeptide of the
CC invention. HHPDZ65 is useful for the treatment of stroke, pain, epilepsy,
CC neurodegenerative diseases and others. The DNAs and proteins are useful
CC in a method for screening to identify compounds which stimulate or
CC inhibit the function of the HHPDZ65 proteins. The polypeptides are useful
CC in a process for diagnosing a disease or a susceptibility to a disease in
CC a subject related to expression or activity of the HHPDZ65 polypeptides.
XX
SQ Sequence 2711 BP; 559 A; 820 C; 852 G; 478 T; 2 other;

Query Match 0.6%; Score 17; DB 19; Length 2711;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2624 cacggtcccccaggagg 2640
||||||||||||||||
Db 1879 cacggtcccccaggagg 1895

RESULT 75
N81166/c
ID N81166 standard; DNA; 2971 BP.
XX
AC N81166;
XX
DT 29-OCT-1990 (first entry)
XX
DE fdhF gene.
XX
KW repression/expression of foreign genes by oxygen/formate; ss.
XX
OS synthetic.
XX
FH Key Location/Qualifiers
FT CDS 749..2896
FT /*tag= a
FT /product=fdhF
FT RBS 737..740
FT /*tag= b
XX
PN EP285152-A.
XX
PD 05-OCT-1988.
XX
PF 30-MAR-1988; 88EP-0105219.
XX
PR 31-MAR-1987; 87DE-3710633.
PR 09-OCT-1987; 87DE-3735381.
XX
PA (BOEF) BOEHRINGER MANNHEIM GMBH.
XX
PI Birkmann A, Bock A;
XX
DR WPI; 1988-294149/42.
XX
PT New recombinant DNA contg consensus sequence and specific promoter -
PT providing foreign gene repression under anaerobic conditions
XX
PS Disclosure; ; 12pp; German.
XX
CC Recombinant DNA contg a consensus sequence common to genes that are
CC repressible by oxygen and inducible by formate under anaerobic
CC conditions and the fdhF promoter used as a simple expression system
CC for foreign genes. Expression does not require temp shifts or addition
CC of inducers. (Formate is produced naturally by microorganisms in the
CC late, anaerobic growth phase).
CC See also N81165.
XX
SQ Sequence 2971 BP; 721 A; 771 C; 844 G; 635 T; 0 other;

Query Match 0.6%; Score 17; DB 9; Length 2971;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 112 ccatatcgcaggcaccc 128
||||||||||||||||
Db 1781 CCATATCGCAGGCACCC 1765

RESULT 76
Q13115
ID Q13115 standard; cDNA; 3000 BP.
XX
AC Q13115;
XX
DT 22-OCT-1991 (first entry)

```
XX
DE   Encodes partial murine Natural Killer receptor.
XX
KW   NK; cytotoxic drugs; tumour cell; immunotherapy; mouse; ss.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   sig_peptide     1..21
FT                   /*tag= a
FT                   /note= "partial"
FT   mat_peptide     22..2946
FT                   /*tag= b
FT                   /product= murine NK receptor
XX
PN   US7535206-A.
XX
PD   09-JUL-1991.
XX
PF   08-JUN-1990;   90US-0143578.
XX
PR   08-JUN-1990;   90US-0535206.
XX
PA   (USSH ) NAT INST OF HEALTH.
XX
PI   Ortaldo J,  Young H,  Anderson S;
XX
DR   WPI; 1991-245694/33.
DR   P-PSDB; R13320.
XX
PT   DNA encoding a natural killer cell receptor - used to develop
PT   prods. for the immuno-detection and immuno-therapy of tumours
XX
PS   Disclosure; Fig 2; 30pp; English.
XX
CC   Overlapping clones which make up this sequence were isolated from
CC   a mouse Peripheral Blood Lymphocyte lambda gt10 cDNA library.
CC   The protein encoded by this sequence is purified and can mediate
CC   the cytolytic activity of mammalian cells. It specifically
CC   distinguishes tumour cells making it a candidate for the development
CC   of products for the immunodetection and immunotherapy of tumours.
CC   See also Q13114.
XX
SQ   Sequence 3000 BP; 997 A; 661 C; 716 G; 626 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 12;  Length 3000;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     684 tcttcagactccgagtc 700
           |||||||||||||||||
Db    2743 tcttcagactccgagtc 2759


RESULT  77
V18471
ID   V18471 standard; cDNA; 3156 BP.
XX
AC   V18471;
XX
DT   14-SEP-1998  (first entry)
XX
DE   T-cell surface antigen CD97 cDNA.
XX
KW   T-cell surface antigen; CD97; human; inflammation; angiogenesis;
KW   atherosclerosis; human; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             49..2556
FT                   /*tag= a
FT   sig_peptide     49..99
FT                   /*tag= b
FT   mat_peptide     100..2553
FT                   /*tag= c
XX
PN   WO9817796-A2.
XX
PD   30-APR-1998.
XX
PF   24-OCT-1997;   97WO-US19772.
XX
PR   25-OCT-1996;   96US-0027871.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI   Kelly K;
XX
DR   WPI; 1998-261492/23.
DR   P-PSDB; W48756.
XX
PT   New soluble CD97 alpha subunit isoform(s) - used to develop
PT   products for the detection and treatment of inflammation,
PT   atherosclerosis and angiogenesis
XX
PS   Disclosure; Fig 1; 101pp; English.
XX
CC   This polynucleotide comprises clone pAT276 that codes for human
CC   T-cell surface antigen CD97 (see W48756).  It was isolated from a
CC   T-cell library enriched for mitogen-induced genes.  The invention
CC   relates to the previously unrecognised alpha subunit of CD97 that
CC   acts in the establishment and maintenance of inflammation.  Soluble
CC   CD97 acts as an adhesion factor for endothelial cells and smooth
CC   muscle cells, implicating it as a modulator of atherosclerosis.
CC   CD97 alpha also acts as a motility factor to cells bearing the
CC   alpha(V)beta-3 receptor, indicative of a role in angiogenesis.
CC   Soluble CD97 alpha1, alpha2, and alpha3 subunits (having different
CC   combinations of EGF repeats) all originate as a proprotein with the
CC   beta subunit (see W48756).  Host cells transfected with a nucleic
CC   acid encoding a CD97 alpha subunits are claimed.  CD97 alpha
CC   subunit polypeptides, nucleic acids, antibodies and antagonists
CC   (e.g. CD97 subunit antisense nucleic acids) are used in claimed
CC   methods for: determining the degree of inflammation at a site:
CC   identifying compounds that inhibit soluble CD97 alpha subunit
CC   expression; inhibiting angiogenesis associated with chronic
CC   inflammation; inhibiting atherosclerosis; and treating or inhibiting
CC   CD97 associated inflammation.
XX
SQ   Sequence 3156 BP; 652 A; 973 C; 860 G; 671 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 19;  Length 3156;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1336 ccaggagggagtggcag 1352
           |||||||||||||||||
Db    2683 ccaggagggagtggcag 2699


RESULT  78
Z27969
ID   Z27969 standard; DNA; 3156 BP.
XX
AC   Z27969;
XX
DT   05-JAN-2000  (first entry)
XX
DE   Human CD97 protein encoding DNA.
XX
KW   Human; 7-transmembrane receptor; lectomedin; lectin-binding; mucin;
KW   olfactomedin; cellular adhesion; atherosclerosis; gene therapy;
KW   vascular disease; CD97; ss.
XX
```

```
OS   Homo sapiens.
XX
PN   WO9945111-A1.
XX
PD   10-SEP-1999.
XX
PF   04-MAR-1999;   99WO-US04676.
XX
PR   04-MAR-1998;   98US-0076782.
XX
PA   (ICOS-) ICOS CORP.
XX
PI   Hayflick JS;
XX
DR   WPI; 1999-571596/48.
DR   P-PSDB; Y41090.
XX
PT   New human lectomedin receptor polypeptide, used to identify specific
PT   binding partners for treating e.g. vascular disease -
XX
PS   Example 1; Page 106-110; 166pp; English.
XX
CC   The invention provides purified and isolated human 7-transmembrane
CC   receptor lectomedin polypeptide or its fragments. The lectomedin
CC   polypeptide comprises extracellular lectin-binding, olfactomedin-like
CC   and mucin-like domains. The polypeptide can be produced by standard
CC   recombinant methodology. The polypeptide is involved in cellular adhesion
CC   and cytoplasmic metabolic pathways that are modulated by extracellular
CC   signaling. Specific binding to lectomedin-1 expressed on smooth muscle
CC   cells may be required for proliferation of these cells in
CC   atherosclerosis. The polypeptide is used to raise specific antibodies,
CC   and to identify specific binding agents that modulate (increase or
CC   decrease) its activity. The lectomedin nucleic acids are used as source
CC   of probes and primers, and of therapeutic antisense, ribozyme or triplex-
CC   forming agents, and in gene therapy to restore deficient lectomedin
CC   activity. Specific binding agents of lectomedin are are used for treating
CC   diseases that involve lectomedin activity, e.g. vascular diseases such as
CC   atherosclerosis. The present sequence represents the DNA encoding the
CC   human CD97 protein.
XX
SQ   Sequence 3156 BP; 651 A; 974 C; 860 G; 671 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 20;  Length 3156;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1336 ccaggagggagtggcag 1352
           |||||||||||||||||
Db    2683 ccaggagggagtggcag 2699


RESULT  79
Q48468/c
ID   Q48468 standard; cDNA; 3457 BP.
XX
AC   Q48468;
XX
DT   21-MAR-1994  (first entry)
XX
DE   Nitrate reductase Nia2 gene derivative.
XX
KW   Nitrate reductase; germination; flowering; ripening; development;
KW   growth stimulation; Agrobacterium tumefaciens; nitrate; ss.
XX
OS   Nicotiana sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             144..2858
FT                   /*tag=  a
FT                   /product= Nitrate reductase.
XX
PN   WO9318154-A.
XX
PD   16-SEP-1993.
XX
PF   05-MAR-1993;   93WO-FR00222.
XX
PR   05-MAR-1992;   92FR-0002658.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX
PI   Caboche M,  Chupeau Y,  Dorlhac F,  Morot-gaudry J;
PI   Vincentz M;
XX
DR   WPI; 1993-303468/38.
DR   P-PSDB; R41757.
XX
PT   Inducing over-expression of nitrate reductase in plants - esp. by
PT   incorporation of foreign gene, for stimulating early development
PT   and reducing nitrate accumulation
XX
PS   Claim 7; Figure 3; 37pp; English.
XX
CC   Overexpression of nitrate reductase (NR) can stimulate the early
CC   development of plants, shortening the duration of the vegetative
CC   phase and causing earlier germination, flowering and ripening by
CC   about two weeks.  Overexpression of NR can also cause the level of
CC   nitrate stored in a plant to be reduced, reducing risks to health
CC   and also possibly improving organoleptic qualities.  The NR gene was
CC   introduced into plants by transforming a strain of Agrobacterium
CC   tumefaciens with a plasmid and using it to infect the subject
CC   plant(s).
XX
SQ   Sequence 3457 BP; 1014 A; 700 C; 752 G; 991 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 14;  Length 3457;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    909 gaaggaagagagatttt 925
          |||||||||||||||||
Db     94 GAAGGAAGAGAGATTTT 78


RESULT  80
T93499
ID   T93499 standard; cDNA; 3796 BP.
XX
AC   T93499;
XX
DT   17-FEB-1998  (first entry)
XX
DE   Xenopus frog protein "chordin" encoding cDNA.
XX
KW   Xenopus protein; chordin; dorsal tissue; neural tissue; vertebrate;
KW   endodermal differentiation; treatment; neurodegenerative disease;
KW   nerve cell; transforming growth factor; TGF; secreted protein; ss.
XX
OS   Xenopus laevis.
XX
FH   Key             Location/Qualifiers
FT   CDS             297..3122
FT                   /*tag=  a
FT                   /product= Xenopus_protein_chordin
FT   sig_peptide     297..353
FT                   /*tag=  b
FT                   /note= "hydrophobic signal peptide"
FT   mat_peptide     354..3179
FT                   /*tag=  c
FT                   /note= "putative secreted protein"
XX
PN   US5679783-A.
XX
PD   21-OCT-1997.
```

XX
RF 22-NOV-1994; 94US-0343760.
XX
AR 22-NOV-1994; 94US-0343760.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI De Robertis EM, Sasai Y;
XX
DR WPI; 1997-525754/48.
DR P-PSDB; W31559.
XX
PT DNA encoding Xenopus frog protein - that induces dorsal and neural
PT development and endodermal differentiation in vertebrates
XX
PS Claim 1; Columns 19-22; 27pp; English.
XX
CC This cDNA encodes a Xenopus protein "chordin". The functional recombinant
CC protein chordin has a defined sequence of 941 amino acids and can induce
CC dorsal and neural development and endodermal differentiation in
CC vertebrates. The presence of a hydrophobic signal sequence, four possible
CC N-glycosylation sites and conserved Cys-rich repeat regions suggest that
CC chordin is a secreted protein. The DNA sequence can be operatively linked
CC with an expression vector, to form a construct and a transformant can be
CC obtained by introducing the construct into a host. Chordin may be useful
CC as a component of culture media for culturing cells such as nerve or
CC muscle cells, for treating neurodegenerative diseases and damaged nerve
CC cells.
XX
SQ Sequence 3796 BP; 1046 A; 841 C; 958 G; 951 T; 0 other;

Query Match 0.6%; Score 17; DB 18; Length 3796;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 270 ctctacgtcttctccga 286
||||||||||||||||
Db 1036 ctctacgtcttctccga 1052

RESULT 81
N70558
ID N70558 standard; DNA; 4216 BP.
XX
AC N70558;
XX
DT 29-APR-1991 (first entry)
XX
DE Sequence of nitrogen fixation gene H (nifH) promoter and coding
DE region and the nifH-nifD intergenic region.
XX
KW Rhizobium expression vector; plant expression vector;
KW Bradyrhizobium; ss.
XX
OS Rhizobium japonicum strain USDA 191.
XX
FH Key Location/Qualifiers
FT promoter 1921..1929
FT /*tag= a
FT promoter 1933..1936
FT /*tag= b
FT RBS 2012..2016
FT /*tag= c
FT CDS 2024..2914
FT /*tag= d
FT /label= nifH
FT CDS 3011..4216
FT /*tag= e
FT /label= nifD
FT misc_feature 1770..2023
FT /*tag= f
FT /note= "claimed"

XX
PN EP211661-A.
XX
PD 25-FEB-1987.
XX
PF 07-AUG-1986; 86EP-0306105.
XX
PR 07-AUG-1985; 85US-0763800.
XX
PA (LUBR) LUBRIZOL GENETICS I.
PA (LUBR-) LUBRIZOL GENETICS I.
XX
PI Appelbaum ER;
XX
DR WPI; 1987-051801/08.
XX
PT Nif promoter of fast-growing Rhizobium japonicum - used to drive
PT transcription in rhizobium of heterologous structural genes
XX
PS Disclosure; Fig 1; 37pp; English.
XX
CC Since the promoter region of the nifH operon has been isolated,
CC characterised and cloned, it is possible to delete the nifH and
CC nifHD nitrogenase genes and replace them with structural genes
CC isolated from an extraneous source. The extraneous genes thus placed
CC under the control of the nifH promoter can then be inserted into a
CC plasmid vector followed by conjugation into a fast-growing R.
CC japonicum strain.
XX
SQ Sequence 4216 BP; 906 A; 1254 C; 1251 G; 793 T; 12 other;

Query Match 0.6%; Score 17; DB 8; Length 4216;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1213 agctcaacctcatccac 1229
||||||||||||||||
Db 3903 agctcaacctcatccac 3919

RESULT 82
T45351
ID T45351 standard; cDNA; 4258 BP.
XX
AC T45351;
XX
DT 18-MAR-1997 (first entry)
XX
DE Human colon carcinoma kinase 4 (CCK-4) cDNA.
XX
KW Colon carcinoma kinase 4; CCK-4; receptor tyrosine kinase;
KW signal transduction; colon cancer; diagnosis; gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 193..3405
FT /*tag= a
FT sig_peptide 193..270
FT /*tag= b
FT mat_peptide 271..3402
FT /*tag= c
XX
PN WO9637610-A2.
XX
PD 28-NOV-1996.
XX
PF 24-MAY-1996; 96WO-IB00696.
XX
PR 25-MAY-1995; 95US-0452630.
XX
PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

```
X
A   Alves F,  Mossie K,  Ullrich A;
X
R   WPI; 1997-021219/02.
R   P-PSDB; W08747.
X
T   New isolated receptor tyrosine kinase, CCK-4 - used for developing
T   prods. for the diagnosis and treatment of CCK-4 signal transduction
T   disorders, partic. colon cancer
X
S   Claim 1; Fig 1a-d; 129pp; English.
X
C   A cDNA sequence (T45351) codes for a novel human receptor tyrosine
C   kinase, colon carcinoma kinase-4 (W08747), or CCK-4, which is
C   preferentially expressed in cancerous colon tissue, compared with
C   normal colon.  The cDNA sequence is a consensus of 6 overlapping
C   cDNA clones isolated from a human placenta library using primers
C   based on conserved motifs of protein tyrosine kinase catalytic
C   domains.  CCK-4 nucleic acids can be used in the prodn. of
C   recombinant CCK-4 polypeptides, and as probes in the diagnosis and
C   screening of CCK-4 signal transduction disorders, esp. colon
C   cancer, and in gene therapy.
X
S   Sequence 4258 BP; 868 A; 1252 C; 1282 G; 856 T; 0 other;


 Query Match             0.6%;  Score 17;  DB 18;  Length 4258;
 Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
 Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1338 aggagggagtggcagag 1354
           |||||||||||||||||
Db   2648 aggagggagtggcagag 2664


RESULT  83
Z51806
ID  Z51806 standard; cDNA; 4308 BP.
XX
AC  Z51806;
XX
DT  04-JUL-2000  (first entry)
XX
DE  Full length expanded cDNA sequence for human BAG-5 protein.
XX
KW  Human BAG-5; Bcl-2 associated athanogene-5; apoptosis; cell migration;
KW  tumour cell metastasis; Hsc70/Hsp70-regulating protein; metastasis;
KW  tumour cell proliferation; steroid hormone receptor function; ss.
XX
OS  Homo sapiens.
XX
FH  Key             Location/Qualifiers
FT  CDS             247..1590
FT                  /*tag= a
FT                  /product= "Human BAG-5 protein"
XX
PN  WO200014106-A1.
XX
PD  16-MAR-2000.
XX
PF  09-SEP-1999;   99WO-US21053.
XX
PR  09-SEP-1998;   98US-0150489.
XX
PA  (BURN-) BURNHAM INST.
XX
AU  Reed JC,  Takayama S;
XX
DR  WPI; 2000-256937/22.
DR  P-PSDB; Y70517.
XX
PT  BAG-1 related proteins from humans, Caenorhabditis elegans and
PT  Schizosaccharomyces pombe useful for modulating tumor cell
PT  proliferation, cell migration and metastasis and steroid hormone
PT  receptor function -
XX
PS  Claim 12; Fig 17; 132pp; English.
XX
CC  The present expanded cDNA sequence encodes human BAG-5 (Bcl-2 associated
CC  athanogene-5) protein. BAG is a Hsc70/Hsp70-regulating protein
CC  (Hsc70/Hsp70 is a molecular chaperone that participates in controlling
CC  protein bioactivity, degradation, complex assembly/disassembly and
CC  translocation across membranes). It competes with Hip for binding to the
CC  Hsc70/Hsp70 ATPase binding domain and promotes substrate release. Gene
CC  transfection studies indicate that BAG proteins influence a wide variety
CC  of cellular phenotypes through their interactions with Hsc70/Hsp70,
CC  including increasing resistance to apoptosis, promoting cell
CC  proliferation, enhancing tumour cell migration and metastasis and
CC  altering transcriptional activity of steroid hormones. BAG also
CC  stimulates Hsc70-mediated adenosine triphosphate (ATP) hydrolysis by
CC  accelerating ADP/ATP exchange.
XX
SQ  Sequence 4308 BP; 1201 A; 897 C; 987 G; 1222 T; 1 other;


 Query Match             0.6%;  Score 17;  DB 21;  Length 4308;
 Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
 Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2126 tttggaagaggaagcag 2142
           |||||||||||||||||
Db   1017 tttggaagaggaagcag 1033


RESULT  84
Q51426
ID  Q51426 standard; cDNA; 4488 BP.
XX
AC  Q51426;
XX
DT  20-MAY-1994  (first entry)
XX
DE  Human FACC cDNA clone #1.
XX
KW  Fanconi Anemia Group C; FACC;  complementing cDNA; varient;
KW  open reading frame; diagnosis; Fanconi anemia; ss.
XX
OS  Homo sapiens.
XX
FH  Key             Location/Qualifiers
FT  CDS             174..1850
FT                  /*tag= a
FT                  /product= Human FACC
FT  polyA_signal    2240..2245
FT                  /*tag= b
FT  polyA_signal    3042..3047
FT                  /*tag= c
FT  repeat_unit     3163..3175
FT                  /*tag= d
FT                  /rpt_type= OTHER
FT                  /note= "palindrome"
FT  repeat_unit     3289..3322
FT                  /*tag= e
FT                  /rpt_type= TANDEM
FT  repeat_unit     3323..4455
FT                  /*tag= f
FT                  /rpt_type= TANDEM
XX
PN  WO9322435-A.
XX
PD  11-NOV-1993.
XX
PF  27-APR-1993;   93WO-CA00178.
XX
PR  29-APR-1992;   92US-0876285.
PR  21-JUL-1992;   92US-0918313.
```

```
-R   15-JAN-1993;   93US-0003963.
XX
-A   (HOSP-) HOSPITAL FOR SICK CHILDREN.
-A   (UNME-) UNITED MEDICAL & DENTAL SCHOOL GUYS.
XX
-I   Buchwald M,  Mathew CG,  Strathdee CA,  Wevrick R;
XX
-R   WPI; 1993-368794/46.
DR   P-PSDB; R44139.
XX
PT   Human cDNA which complements Fanconi Anaemia gp. C - used to
PT   develop prods. for use in diagnosis, study and therapy of Fanconi
PT   Anaemia
XX
PS   Claim 1; Page 97-101; 137pp; English.
XX
CC   The sequences given in Q51426-28 represent cDNA varients from the
CC   Fanconi Anemia Group C Complementing (FACC) cDNA.  These three cDNA
CC   molecules are cellular varients of a single cDNA transcribed from the
CC   same gene.  The three cDNAs each contain an identical open reading
CC   frame encoding the FACC protein.  The FACC protein may be used for
CC   the diagnosis and study of Fanconi anemia.
XX
SQ   Sequence 4488 BP; 1052 A; 1092 C; 1168 G; 1176 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 14;  Length 4488;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2105 tgaagccaccctggaag 2121
           |||||||||||||||||
Db   4189 tgaagccaccctggaag 4205


RESULT  85
V33945
ID   V33945 standard; cDNA; 4567 BP.
XX
AC   V33945;
XX
DT   15-FEB-1999  (first entry)
XX
DE   Fanconi anaemia complementation group C (FAC) cDNA.
XX
KW   Fanconi anaemia complementation group C; FAC; apoptosis;
KW   haematopoiesis; bone marrow; chemotherapy; gene therapy; human; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             256..1929
FT                   /*tag=  a
XX
PN   WO9851792-A1.
XX
PD   19-NOV-1998.
XX
PF   15-MAY-1998;   98WO-US09975.
XX
PR   15-MAY-1997;   97US-0046546.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
XX
PI   Youssoufian H;
XX
DR   WPI; 1999-009774/01.
DR   P-PSDB; W68546.
XX
PT   New conjugate of Fanconi anaemia molecule and peptide selective for
PT   haematopoietic precursor cells - inhibits apoptosis of these cells,
PT   for treating Fanconi anaemia and patients undergoing high-dose
PT   chemotherapy for cancer
XX
PS   Claim 6; Page 40-45; 72pp; English.
XX
CC   This cDNA clone includes a coding region for human Fanconi anaemia
CC   complementation group C (FAC, see W68546), a protein that modulates
CC   apoptosis in haematopoietic progenitor cells (HPC).  The invention
CC   provides conjugates, including fusion proteins, comprising FAC
CC   and a targetting molecule which binds to a cell surface protein of
CC   the HPC and is internalised.  Such targetting molecules include
CC   interleukin-3 (see W68547) and antibodies which recognise CD33 (see
CC   W68548-49).  The conjugate, or a nucleic acid encoding it, can be
CC   used to deliver FAC to an HPC, specifically to inhibit apoptosis,
CC   particularly in patients exposed to high doses of chemotherapy for
CC   treatment of non-myeloid cancers, also to treat Fanconi anaemia (by
CC   complementation of the genetic defect).  Treatment of HPC is done
CC   in vitro, ex vivo (e.g. for recombinant production of conjugate in
CC   cell cultures) or in vivo.  Treatment with FAC may eliminate the
CC   need for extensive bone marrow transplants to restore
CC   haematopoiesis after chemotherapy.
XX
SQ   Sequence 4567 BP; 1080 A; 1135 C; 1177 G; 1175 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 20;  Length 4567;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2105 tgaagccaccctggaag 2121
           |||||||||||||||||
Db   4268 tgaagccaccctggaag 4284


RESULT  86
Q32352
ID   Q32352 standard; DNA; 5674 BP.
XX
AC   Q32352;
XX
DT   22-APR-1993  (first entry)
XX
DE   MAGE-1 nucleic acid.
XX
KW   melanoma antigen; MAGE TRA; melanoma antigen tumor rejection antigen;
KW   tumor rejection antigen precursor; MAGE; antigen E; gene family; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             3881..4711
FT                   /*tag=  a
XX
PN   WO9220356-A.
XX
PD   26-NOV-1992.
XX
PF   22-MAY-1992;   92WO-US04354.
XX
PR   23-MAY-1991;   91US-0705702.
PR   09-JUL-1991;   91US-0728838.
PR   23-SEP-1991;   91US-0764364.
PR   12-DEC-1991;   91US-0807043.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
XX
PI   Boon T,  Chomez P,  De Plaen E,  Lurquin C,  Traversari C;
PI   Van Den Eynde B,  Van Der Bruggen P,  Van Pel A;
XX
DR   WPI; 1992-415460/50.
XX
PT   Nucleic acid mol. encoding a human tumour rejection antigen
PT   precursor - useful as an immunostimulant in a vaccine for
PT   treating and preventing cancers, also useful in diagnosis
XX
```

```
S   Disclosure; Page 71-73; 142pp; English.
X
C   The sequences given in Q32352-69 represent a new family of genes
C   refered to as melanoma antigens (MAGE).  The cDNAs of this gene
C   family were identified during the isolation of the antigen E gene.
C   The MAGE cDNAs, when tested, did not transfer expression of antigen
C   E, but they did show substantial homology to the antigen E cDNA
C   sequence.  The MAGE DNAs share a certain degree of homology with each
C   other and are expressed in tumour cells including several types of
C   human tumor cells as well as in human tumors.  MAGE expression is not
C   restricted to melanomas.  MAGE refers to a family of tumor rejection
C   antigen precursors.  The antigens resulting from these genes are
C   referred to as MAGE TRAs or melanoma antigen tumor rejection antigens.
C   See also Q32351.
XX
Q   Sequence 5674 BP; 1277 A; 1644 C; 1568 G; 1185 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 13;  Length 5674;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

y   1652 ggtcctgggcaccctgg 1668
        |||||||||||||||||
b   3994 ggtcctgggcaccctgg 4010


ESULT  87
Q72477
D   Q72477 standard; DNA; 5674 BP.
X
C   Q72477;
XX
T   22-JUN-1995  (first entry)
X
S   Tumour rejection antigen MAGE-1 encoding DNA.
X
W   Tumour rejection antigen; melanoma antigen-1; MAGE-1; MAGE-3;
W   cancer; cytolytic T cells; antigen D; human leucocyte antigen;
W   ss.
X
S   Homo sapiens.

H   Key             Location/Qualifiers
    CDS             3881..4711
                    /*tag=  a
X
N   WO9423031-A.
X
D   13-OCT-1994.
X
    17-MAR-1994;   94WO-US02877.
X
R   26-MAR-1993;   93US-0037230.
X
A   (LUDW-) LUDWIG INST CANCER RES.
X
I   Boon-falleur T,  Gaugler B,  Van DEN EYNDE B,  Van DER BRUGGEN P;
XX
R   WPI; 1994-333192/41.
X
T   New tumour rejection antigen precursor MAGE3 - useful in
    treatment and diagnosis of cancer
X
S   Example 26; Page 59; 105pp; English.
X
C   Q72477 is the DNA sequence which encodes melanoma antigen-1
C   (MAGE-1). Another melanoma antigen MAGE-3 is encoded by Q72470,
C   this is a tumour rejection antigen precursor. Melanomas
C   characterised by the expression of MAGE-3 can be detected, or
C   monitored, by contacting a test sample with an agent that can
C   recognise MAGE-3. The melanoma can be treated by the administration
C   of cytolytic T cells specific for the complex of antigen D (the
CC  mature rejection antigen derived from MAGE-3) and a human leucocyte
CC  antigen (esp. HLA-A1).
XX
SQ  Sequence 5674 BP; 1276 A; 1644 C; 1569 G; 1185 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 15;  Length 5674;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy  1652 ggtcctgggcaccctgg 1668
         |||||||||||||||||
Db  3994 ggtcctgggcaccctgg 4010


RESULT  88
X84113
ID  X84113 standard; DNA; 5674 BP.
XX
AC  X84113;
XX
DT  08-SEP-1999  (first entry)
XX
DE  MAGE-1 gene.
XX
KW  Tumour rejection antigen; vaccine; cancer; MAGE-1 gene; ss.
XX
OS  Homo sapiens.
XX
PN  US5925729-A.
XX
PD  20-JUL-1999.
XX
PF  02-MAY-1994;   94US-0142368.
XX
PR  02-MAY-1994;   94US-0142368.
PR  23-MAY-1991;   91US-0705702.
PR  09-JUL-1991;   91US-0728838.
PR  23-SEP-1991;   91US-0764365.
PR  12-DEC-1991;   91US-0807043.
XX
PA  (LUDW-) LUDWIG INST CANCER RES.
XX
PI  Boon T,  Chomez P,  De Plaen E,  Lurquin C,  Traversari C;
PI  Van Den Eynde B, Van Der Bruggen P,  Van Pel A;
XX
DR  WPI; 1999-418294/35.
XX
PT  New tumour rejection antigen is useful as a vaccine against
PT  cancerous diseases
XX
PS  Disclosure; Column 39-46; 58pp; English.
XX
CC  This sequence represents the MAGE-1 gene sequence.
CC  The invention relates to a tumour rejection antigen sequence that is
CC  useful as a tumour rejection antigen for vaccination against cancerous
CC  conditions.
XX
SQ  Sequence 5674 BP; 1276 A; 1644 C; 1569 G; 1185 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 20;  Length 5674;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy  1652 ggtcctgggcaccctgg 1668
         |||||||||||||||||
Db  3994 ggtcctgggcaccctgg 4010


RESULT  89
T42117
ID  T42117 standard; cDNA; 5720 BP.
```

```
XX
AC   T42117;
XX
DT   22-JAN-1997  (first entry)
XX
DE   Lats gene encoding large tumour suppressor.
XX
KW   Lats gene; large tumour suppressor; Drosophila melanogaster;
KW   fruitfly; polyadenylation site; protein-serine/threonine-kinase;
KW   cell proliferation; antisense; dominant-negative; cancer;
KW   degenerative disorder; trauma; growth deficiency; therapy;
KW   antitumour; vulnerary; diagnostic; transgenic plant;
KW   transgenic animal; growth; senescence; ds.
XX
OS   Drosophila melanogaster.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1..191
FT                   /*tag= a
FT                   /note= "Sequence from clone cDNA-9"
FT   5'UTR           1..1102
FT                   /*tag= b
FT   conflict        81
FT                   /*tag= c
FT                   /note= "A in genomic sequence"
FT   misc_feature    192..5720
FT                   /*tag= d
FT                   /note= "Sequence from clone cDNA-A2"
FT   misc_feature    302..303
FT                   /*tag= e
FT                   /note= "Intron-1 splice site"
FT   conflict        330
FT                   /*tag= f
FT                   /note= "T in genomic sequence"
FT   conflict        606..608
FT                   /*tag= g
FT                   /note= "GGA in genomic sequence"
FT   conflict        821
FT                   /*tag= h
FT                   /note= "G in genomic sequence"
FT   conflict        834
FT                   /*tag= i
FT                   /note= "A in genomic sequence"
FT   CDS             1103..4402
FT                   /*tag= j
FT                   /product= Large tumour suppressor (lats protein)
FT   misc_feature    1225..1226
FT                   /*tag= k
FT                   /note= "Intron-2 splice site"
FT   mutation        1371..1818
FT                   /*tag= l
FT                   /note= "lats-al deletion"
FT   conflict        1447
FT                   /*tag= m
FT                   /note= "C in genomic sequence"
FT   conflict        1719
FT                   /*tag= n
FT                   /note= "G in genomic sequence"
FT   conflict        1939
FT                   /*tag= o
FT                   /note= "T in genomic sequence"
FT   conflict        2137
FT                   /*tag= p
FT                   /note= "C in genomic sequence"
FT   conflict        2431
FT                   /*tag= q
FT                   /note= "C in genomic sequence"
FT   conflict        2458
FT                   /*tag= r
FT                   /note= "T in genomic sequence"
FT   conflict        2551
FT                   /*tag= s
FT                   /note= "G in genomic sequence"
FT   conflict        2554
FT                   /*tag= t
FT                   /note= "G in genomic sequence"
FT   conflict        2555
FT                   /*tag= u
FT                   /note= "AAC in genomic sequence"
FT   conflict        2556
FT                   /*tag= v
FT                   /note= "A in genomic sequence"
FT   conflict        2593
FT                   /*tag= w
FT                   /note= "G in genomic sequence"
FT   conflict        2646
FT                   /*tag= x
FT                   /note= "GGGAGCGGGATCAAC insertion in genomic DNA"
FT   misc_feature    3488..3489
FT                   /*tag= y
FT                   /note= "Intron-3 splice site"
FT   misc_feature    3609..3610
FT                   /*tag= z
FT                   /note= "Intron-4 splice site"
FT   misc_feature    3725..3726
FT                   /*tag= aa
FT                   /note= "Intron-5 splice site"
FT   misc_feature    3804..3805
FT                   /*tag= ab
FT                   /note= "Intron-6 splice site"
FT   conflict        3889
FT                   /*tag= ac
FT                   /note= "A in genomic sequence"
FT   conflict        3937
FT                   /*tag= ad
FT                   /note= "C in genomic sequence"
FT   conflict        3940
FT                   /*tag= ae
FT                   /note= "T in genomic sequence"
FT   misc_feature    4000..4001
FT                   /*tag= af
FT                   /note= "Intron-7 splice site"
FT   conflict        4021
FT                   /*tag= ag
FT                   /note= "T in genomic sequence"
FT   conflict        4021
FT                   /*tag= ah
FT                   /note= "C in genomic sequence"
FT   conflict        4210
FT                   /*tag= ai
FT                   /note= "G in genomic sequence"
FT   conflict        4300
FT                   /*tag= aj
FT                   /note= "C in genomic sequence"
FT   conflict        4309
FT                   /*tag= ak
FT                   /note= "C in genomic sequence"
FT   3'UTR           4402..5720
FT                   /*tag= al
FT   polyA_signal    4654..4660
FT                   /*tag= am
FT   conflict        4797
FT                   /*tag= an
FT                   /note= "G in genomic sequence"
FT   conflict        4828..4837
FT                   /*tag= ao
FT                   /note= "Deleted in genomic DNA"
FT   conflict        4871
FT                   /*tag= ap
FT                   /note= "C in genomic sequence"
FT   conflict        4974
FT                   /*tag= aq
FT                   /note= "T in genomic sequence"
FT   misc_feature    5013..5153
FT                   /*tag= ar
FT                   /note= "Region identical to plc-21 transcript"
```

```
FT   conflict        5101
                     /*tag= as
                     /note= "C in genomic sequence"
FT   conflict        5104..5114
                     /*tag= at
                     /note= "TGTAATTAGTG in genomic sequence"
FT   conflict        5143
                     /*tag= au
                     /note= "A in genomic sequence"
FT   conflict        5147
                     /*tag= av
                     /note= "A in genomic sequence"
FT   conflict        5151..5169
                     /*tag= aw
                     /note= "GTGGCCCCCCTCCCTCCCTCCTCAT in genomic
                           sequence"
FT   conflict        5247..5253
                     /*tag= ax
                     /note= "Deleted in genomic DNA"
FT   conflict        5646..5663
                     /*tag= ay
                     /note= "AAAAGCAAATTAATAAAT in genomic sequence"
FT   polyA_signal    5657..5663
                     /*tag= az
XX
WO9630402-A1.
XX
03-OCT-1996.
XX
26-MAR-1996;   96WO-US04101.
XX
27-MAR-1995;   95US-0411111.
XX
(UYYA ) UNIV YALE.
XX
Tao W,  Wang W,  Xu T,  Yu W,  Zhang S;
XX
WPI; 1996-455275/45.
P-PSDB; W05177.
XX
New isolated large tumour suppressor gene - used to develop prods.
for inhibiting cell proliferation or for enhancing proliferation
XX
Disclosure; Page 109-115; 215pp; English.
XX
This sequence encodes the Drosophila melanogaster large tumour
suppressor lats protein, and is a composite of 2 cDNAs (an initial
fragment from cDNA-9 and the rest from cDNA-A2).  The sequence has
been isolated from a total imaginal disc cDNA library.  The
corresponding genomic sequence is identical, except for 34 minor
differences, and has 7 introns.  Two consensus polyadenylation sites
are present.  A 141-bp sequence at the 3'-end of the lats transcript
is identical to the 5'-end of the untranslated sequence of the

  Query Match              0.6%;  Score 17;  DB 17;  Length 5720;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    814 tggtgctcaaagcaaag 830
          |||||||||||||||||
Db   4820 tggtgctcaaagcaaag 4836


RESULT  90
Z51508
ID   Z51508 standard; DNA; 5720 BP.
XX
AC   Z51508;
XX
DT   21-JUN-2000  (first entry)
XX
DE   Drosophila melanogaster Lats (large tumour suppressor) DNA.
XX
```

```
KW   Fruit fly; Lats; large tumour suppressor; cytostatic; vulnerary;
KW   cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase;
KW   treatment; prevention; screening; cancer; skin; ovarian tumour;
KW   soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;
KW   LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;
KW   dysplasia; degenerative disorder; growth deficiency; physical trauma;
KW   hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.
XX
OS   Drosophila melanogaster.
XX
FH   Key             Location/Qualifiers
FT   CDS             1103..4402
FT                   /*tag= a
FT                   /product= "Lats protein"
FT   polyA_signal    4655..4660
FT                   /*tag= b
FT   misc_feature    5013..5142
FT                   /*tag= c
FT                   /note= "This region is identical to the 1-141
FT                   nucleotides of Drosophila plc-21 transcript"
XX
PN   WO200010602-A1.
XX
PD   02-MAR-2000.
XX
PF   18-AUG-1999;   99WO-US19068.
XX
PR   18-AUG-1998;   98US-0096996.
PR   18-AUG-1998;   98US-0096997.
XX
PA   (UYYA ) UNIV YALE.
XX
PI   Xu T,  Tao W,  St John MAR,  Fei X,  Fukumoto RK,  Zhang S;
PI   Turenchalk GS,  Stewart RA;
XX
DR   WPI; 2000-246496/21.
DR   P-PSDB; Y70393.
XX
PT   Use of lats proteins, complexes of lats and cdc2 for treating cancer
PT   that is refractory to treatment by standard chemotherapy and radiation
PT   therapy, and disorders associated with aberrant levels of cdc2 activity
PT   -
XX
PS   Claim 44; Fig 15; 134pp; English.
XX
CC   The present sequence is a DNA encoding Drosophila Lats (large tumour
CC   suppressor) protein which is a cell overproliferation inhibitor and a
CC   negative regulator of cell cycle-dependent kinase cdc2/cyclin A.
CC   The present sequence is useful for treating cancer that is refractory
CC   to standard chemotherapy or radiation therapy such as hyperplasia,
CC   metaplasia, or dysplasia, and disorders associated with aberrant
CC   levels of cdc2 activity. Conditions treated by promoting cdc2 function
CC   include degenerative disorders, growth deficiencies, hypoproliferative
CC   disorders, physical trauma, lesions, and wounds. An animal model
CC   preferably a mouse, in which a lats gene has been disrupted by homologous
CC   recombination, e.g. a lats knock-out mouse, is used for screening
CC   compounds that can be used to treat or prevent cancer, particularly
CC   skin cancer, soft tissue sarcomas and ovarian tumours, and disorders
CC   associated with pituitary dysfunction e.g. luteinizing hormone (LH)
CC   hypogonadotropic hypogonadism. The lats DNA is also used in gene therapy.
XX
SQ   Sequence 5720 BP; 1684 A; 1491 C; 1457 G; 1088 T; 0 other;


  Query Match              0.6%;  Score 17;  DB 21;  Length 5720;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    814 tggtgctcaaagcaaag 830
          |||||||||||||||||
Db   4820 tggtgctcaaagcaaag 4836
```

```
RESULT  91
Q98902
ID   Q98902 standard; DNA; 5724 BP.
XX
AC   Q98902;
XX
DT   28-FEB-1996  (first entry)
XX
DE   Tumour rejection antigen (MAGE-1) gene.
XX
KW   Tumour rejection antigen; MAGE-1; monoclonal antibody; MAb;
KW   diagnosis; immunoassay; cancer; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             3881..4711
FT                   /*tag= a
FT                   /product= Tumour rejection antigen MAGE-1.
FT                   /note= "The CDS is not indicated in the text of the
FT                          specification but is suggested in the layout
FT                          of the sequence."
XX
PN   WO9520974-A1.
XX
PD   10-AUG-1995.
XX
PF   05-JAN-1995;   95WO-US00095.
XX
PR   01-FEB-1994;   94US-0190411.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
PA   (SLOK ) SLOAN KETTERING INST CANCER RES.
PA   (SLOK ) MEMORIAL SLOAN-KETTERING CANCER CENT.
XX
PI   Boon-falleur T,  Chen Y,  Garin-chesa P,  Old LJ;  Rettig WJ;
PI   Stockert E,  Van der bruggen P;
XX
DR   WPI; 1995-283606/37.
XX
PT   New monoclonal antibody binding specifically to MAGE-1 - useful for
PT   diagnosis and monitoring of cancer, also new hybridomas, recombinant
PT   MAGE-1 and immunogenic peptide(s)
XX
PS   Disclosure; Page 16-19; 33pp; English.
XX
CC   A monoclonal antibody directed against the tumour rejection antigen
CC   (MAGE-1) can be used to detect MAGE-1 in samples by standard
CC   immunoassay methods for diagnosis and monitoring of cancer etc. The
CC   monoclonal antibody is designated MA454 and is produced by the
CC   hybridoma deposited as ATCC HBB11540.  The monoclonal antibody is
CC   specific for MAGE-1, having no reactivity for MAGE-2 or MAGE-3.
CC   Peptide fragments of MAGE-1 (See R80618-20) may be useful as
CC   immunogens for production of the monoclonal antibody and antisera.
XX
SQ   Sequence 5724 BP; 1282 A; 1653 C; 1589 G; 1200 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 16;  Length 5724;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1652 ggtcctgggcaccctgg 1668
           |||||||||||||||||
Db   3994 ggtcctgggcaccctgg 4010


RESULT  92
Q48772
ID   Q48772 standard; DNA; 6464 BP.
XX
AC   Q48772;
XX
DT   21-APR-1994  (first entry)
XX
DE   Arylamine N-acetyl-transferase type 3.
XX
KW   Arylamine N-acetyl-transferase; NAT; polymorphism; diagnosis;
KW   amino; aromatic substance; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             723..1595
FT                   /*tag= a
FT   exon            717..1936
FT                   /*tag= b
FT   polyA_signal    1794..1799
FT                   /*tag= c
FT   polyA_signal    1800..1805
FT                   /*tag= d
XX
PN   EP562547-A.
XX
PD   29-SEP-1993.
XX
PF   23-MAR-1993;   93EP-0104753.
XX
PR   23-MAR-1992;   92JP-0064669.
XX
PA   (SAKA ) OTSUKA PHARM CO LTD.
PA   (TOKM-) TOKYO METROPOLITAN INST. NEUROSCIENCES.
PA   (TOKY-) ZH TOKYOTO SHINKEIKAGAKA SOGO KENKYUSHO.
XX
PI   Deguchi T,  Katsuragi K,  Kinoshita M,  Shin S;
XX
DR   WPI; 1993-304947/39.
DR   P-PSDB; R41246.
XX
PT   Polymorphic human arylamine N-acetyl-transferase genes - used to
PT   diagnose adverse effects caused by amino-contg. aromatic
PT   substances
XX
PS   Claim 6; Page 37-41; 54pp; English.
XX
CC   Sequences of three polymorphic human NAT genes are given in Q48767-
CC   Q48772.  Detection of polymorphic human NAT genes allows diagnosis
CC   of adverse effects to be caused by amino-contg. aromatic substances.
XX
SQ   Sequence 6464 BP; 2087 A; 1189 C; 1282 G; 1906 T; 0 other;


  Query Match             0.6%;  Score 17;  DB 14;  Length 6464;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1032 aggtaccaaggaaaggc 1048
           |||||||||||||||||
Db   5916 aggtaccaaggaaaggc 5932


RESULT  93
V52165
ID   V52165 standard; DNA; 10240 BP.
XX
AC   V52165;
XX
DT   23-OCT-1998  (first entry)
XX
DE   Streptococcus pneumoniae genome fragment SEQ ID NO:32.
XX
KW   Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
KW   computer readable medium; vaccine; pharmaceutical composition; ds.
XX
OS   Streptococcus pneumoniae.
XX
```

```
PN   WO9818931-A2.
XX
PD   07-MAY-1998.
XX
PF   30-OCT-1997;   97WO-US19588.
XX
PR   31-OCT-1996;   96US-0029960.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Barash SC,  Choi GH,  Dillon PJ,  Dougherty BA,  Fannon M;
PI   Kunsch CA,  Rosen CA;
XX
DR   WPI; 1998-272225/24.
XX
PT   Computer-readable medium with recorded Streptococcus pneumoniae
PT   polynucleotide sequences - useful in diagnostic kits and assays, and
PT   pharmaceutical compositions and vaccines for Streptococcus
PT   pneumoniae
XX
PS   Claim 1; Page 326-332; 1409pp; English.
XX
CC   The present invention describes a computer readable medium which has
CC   the nucleotide sequences SEQ ID NO:1 to 391 (V52134 to V52524) recorded
CC   on it, or a representative fragment or a sequence at least 95% identical
CC   to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in SEQ ID NO:1
CC   to 391 (V52134 to V52524) are genomic fragments from Streptococcus
CC   pneumoniae. The present invention also describes an isolated nucleic acid
CC   molecule encoding a homologue of any of the fragments of the S.pneumoniae
CC   genome (SEQ ID NO:1 to 391) where the nucleic acid molecule is produced
CC   by a process comprising: (a) screening a genomic DNA library using as a
CC   probe a target sequence defined by any of the sequences in SEQ ID NO:1
CC   to 391, identifying members of the library which contain sequences
CC   that hybridise to the target sequence and isolating the nucleic acid
CC   molecules from the members; or (b) isolating mRNA, DNA or cDNA produced
CC   from an organism, amplifying nucleic acid molecules whose nucleotide
CC   sequence is homologous to amplification primers derived from the
CC   fragment of the S. pneumoniae genome to prime the amplification and
CC   isolating the amplified sequences. The computer readable medium can be
CC   used in a computer-based system for identifying fragments of the
CC   S. pneumoniae genome of commercial importance, or expression modulating
CC   fragments of the S. pneumoniae genome. Products from the present
CC   invention can be used in diagnosis kits and assays, and pharmaceutical
CC   compositions and vaccines for S. pneumoniae.
XX
SQ   Sequence 10240 BP; 2865 A; 1914 C; 2390 G; 3068 T; 3 other;


  Query Match             0.6%;  Score 17;  DB 19;  Length 10240;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1947 gatttggaagagtttca 1963
           |||||||||||||||||
Db    7643 gatttggaagagtttca 7659


RESULT  94
X20553
ID   X20553 standard; DNA; 10461 BP.
XX
AC   X20553;
XX
DT   05-MAY-1999  (first entry)
XX
DE   Polynucleotide sequence from the genome of Treponema pallidum.
XX
KW   Treponema pallidum infection; syphilis; Borrelia infection; animal;
KW   enzyme production; ds.
XX
OS   Treponema pallidum.
XX
PN   WO9859034-A2.
XX
PD   30-DEC-1998.
XX
PF   23-JUN-1998;   98WO-US13041.
XX
PR   24-JUN-1997;   97US-0050667.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Fraser CM;
XX
DR   WPI; 1999-081273/07.
XX
PT   New isolated Treponema pallidum nucleic acids - used to develop
PT   products for the detection, diagnosis, characterisation, prevention
PT   and therapy of T. pallidum infections, particularly syphilis
XX
PS   Claim 1; Page 491-497; 1150pp; English.
XX
CC   X20500-21243 represent polynucleotide sequences from the genome of
CC   Treponema pallidum. The sequences can be used for detection,
CC   diagnosis, characterisation, prevention and therapy for T. pallidum
CC   infections, particularly syphilis. They can also be used for detecting
CC   diseases related to Borrelia infections in animals, and for the
CC   production of biosynthetic products such as enzymes.
XX
SQ   Sequence 10461 BP; 2325 A; 3411 C; 2493 G; 2213 T; 19 other;


  Query Match             0.6%;  Score 17;  DB 20;  Length 10461;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     161 gcggcacctgcgcacgc 177
           |||||||||||||||||
Db    6768 gcggcacctgcgcacgc 6784


RESULT  95
V74675
ID   V74675 standard; DNA; 10813 BP.
XX
AC   V74675;
XX
DT   16-MAR-1999 .(first entry)
XX
DE   Staphylococcus aureus contig SEQ ID #364.
XX
KW   Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW   cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW   skin infection; surgical wound infection; scalded skin syndrome;
KW   toxic shock syndrome; ds.
XX
OS   Staphylococcus aureus.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    301..360
FT                   /*tag= a
FT                   /note= "these bases represent a line of missing text in
FT                          the sequence listing in the specification. They
FT                          are included to maintain the nucleotide numbering
FT                          given in the specification for this DNA sequence"
FT   misc_feature    2101..2160
FT                   /*tag= b
FT                   /note= "these bases represent a line of missing text in
FT                          the sequence listing in the specification. They
FT                          are included to maintain the nucleotide numbering
FT                          given in the specification for this DNA sequence"
FT   misc_feature    3901..3960
FT                   /*tag= c
FT                   /note= "these bases represent a line of missing text in
FT                          the sequence listing in the specification. They
FT                          are included to maintain the nucleotide numbering
```

```
FT                      given in the specification for this DNA sequence"
FT   misc_feature    5701..5760
FT                   /*tag= d
FT                   /note= "these bases represent a line of missing text in
FT                          the sequence listing in the specification. They
FT                          are included to maintain the nucleotide numbering
FT                          given in the specification for this DNA sequence"
FT   misc_feature    7501..7560
FT                   /*tag= e
FT                   /note= "these bases represent a line of missing text in
FT                          the sequence listing in the specification. They
FT                          are included to maintain the nucleotide numbering
FT                          given in the specification for this DNA sequence"
FT   misc_feature    9301..9360
FT                   /*tag= f
FT                   /note= "these bases represent a line of missing text in
FT                          the sequence listing in the specification. They
FT                          are included to maintain the nucleotide numbering
FT                          given in the specification for this DNA sequence"
XX
PN   EP786519-A2.
XX
PD   30-JUL-1997.
XX
PF   07-JAN-1997;   97EP-0100117.
XX
PR   05-JAN-1996;   96US-0009861.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Barash SC,  Choi GH,  Dillon PJ,  Fannon MR,  Kunsch CA;
PI   Rosen CA;
XX
DR   WPI; 1997-374922/35.
XX
PT   Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT   stored on computer readable medium and used in the production of
PT   anti-S.aureus vaccines
XX
PS   Claim 1; Page 1248-1254; 3271pp; English.
XX
CC   This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC   of the invention. The DNA sequences are recorded on a computer readable
CC   medium, preferably selected from a floppy or hard disk, random access
CC   memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC   the S.aureus DNA sequences allows putative functions to be assigned so
CC   that protein-encoding or regulatory regions of commercial, therapeutic or
CC   industrial importance can be obtained. Specifically, sequences which are
CC   likely to encode antigens have been identified and these polypeptides can
CC   be used in a vaccine composition against S.aureus infection. The
CC   polypeptides can also be used in a kit for the immunodetection of
CC   S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC   including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC   skin and surgical wound infections, scalded skin syndrome, toxic shock
CC   syndrome, etc. Organisms transformed with the DNA sequences can be used
CC   for recombinant production of the polypeptides. The new DNA sequences
CC   (and their fragments) are useful as primers or probes for isolating
CC   homologues of any of the S.aureus DNA sequences contained on the
CC   computer readable medium.
XX
SQ   Sequence 10813 BP; 3513 A; 1895 C; 1594 G; 3451 T; 360 other;


 Query Match             0.6%;  Score 17;  DB 18;  Length 10813;
 Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
 Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    474 ccacaactggaaaaata 490
          |||||||||||||||||
Db  10019 ccacaactggaaaaata 10035


RESULT  96
Q70447
ID   Q70447 standard; DNA; 11236 BP.
XX
AC   Q70447;
XX
DT   21-MAR-1995  (first entry)
XX
DE   Rat nestin gene - its product is useful to identify brain tumours.
XX
KW   nestin gene; brain tumour; neoplastic cells; glial; neuronal;
KW   muscle; neural multipotential stem cell; mammalian brain; detection;
KW   diagnosis; medulloblastoma; gliablastoma; oligodendroglioma; ds.
XX
OS   Rattus rattus.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    2463
FT                   /*tag= a
FT                   /note= "start of primary transcript"
FT   CDS             2589..10821
FT                   /*tag= b
FT                   /product= Nestin protein
FT   misc_feature    3089..3090
FT                   /*tag= c
FT                   /transl_except= pos:3087..3091, aa:His, Arg
FT                   /note= "sequence should be CAY CGG, ie. Y or C has
FT                          been deleted in the sequence given"
FT   misc_feature    3104
FT                   /*tag= d
FT                   /note= "apparent inclusion of a nucleotide"
FT   intron          3375..4339
FT                   /*tag= e
FT   exon            4340..4464
FT                   /*tag= f
FT   intron          4465..6119
FT                   /*tag= g
FT   exon            6120..6193
FT                   /*tag= h
FT   intron          6194..6388
FT                   /*tag= i
FT   exon            6389..10821
FT                   /*tag= j
XX
PN   US5338839-A.
XX
PD   16-AUG-1994.
XX
PF   12-APR-1988;   88US-0180548.
XX
PR   12-APR-1988;   88US-0180548.
PR   02-JUN-1988;   88US-0201762.
PR   25-OCT-1990;   90US-0603803.
PR   22-FEB-1991;   91US-0660412.
PR   19-MAR-1992;   92US-0853913.
XX
PA   (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
PI   Lendahl U,  Mckay RDG;
XX
DR   WPI; 1994-263332/32.
DR   P-PSDB; R60126.
XX
PT   Nucleotide and protein sequences for human and rat nestin -
PT   distinguishes neural multipotential stem cells and brain tumour
PT   cells from more differentiated cell types; for use in the
PT   diagnosis of brain tumours
XX
PS   Claim 1; Column 25-34; 45pp; English.
XX
CC   Q70447 is the rat nestin gene encoding nestin protein (R60126).
CC   Nestin protein expression distinguishes neural multipotential stem
CC   cells and brain tumour cells from the more differentiated neural
CC   cell types (eg., neuronal, glial and muscle cells of the adult brain).
```

CC The nestin protein can be used in diagnosing tumours of the brain,
CC such as medulloblastomas, gliablastomas and oligodendroglioma.
CC (See also Q70448).
XX
SQ Sequence 11236 BP; 2876 A; 2678 C; 3258 G; 2424 T; 0 other;

Query Match 0.6%; Score 17; DB 15; Length 11236;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1877 ccttcaggaaggggctg 1893
|||||||||||||||||
Db 3466 ccttcaggaaggggctg 3482

RESULT 97
Z20535/c
ID X20535 standard; DNA; 21170 BP.
XX
AC X20535;
XX
DT 05-MAY-1999 (first entry)
XX
DE Polynucleotide sequence from the genome of Treponema pallidum.
XX
KW Treponema pallidum infection; syphilis; Borrelia infection; animal;
KW enzyme production; ds.
XX
OS Treponema pallidum.
XX
PN WO9859034-A2.
XX
PD 30-DEC-1998.
XX
PF 23-JUN-1998; 98WO-US13041.
XX
PR 24-JUN-1997; 97US-0050667.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Fraser CM;
XX
DR WPI; 1999-081273/07.
XX
PT New isolated Treponema pallidum nucleic acids - used to develop
PT products for the detection, diagnosis, characterisation, prevention
PT and therapy of T. pallidum infections, particularly syphilis
XX
PS Claim 1; Page 389-401; 1150pp; English.
XX
CC X20500-21243 represent polynucleotide sequences from the genome of
CC Treponema pallidum. The sequences can be used for detection,
CC diagnosis, characterisation, prevention and therapy for T. pallidum
CC infections, particularly syphilis. They can also be used for detecting
CC diseases related to Borrelia infections in animals, and for the
CC production of biosynthetic products such as enzymes.
XX
SQ Sequence 21170 BP; 4629 A; 5015 C; 6107 G; 5390 T; 29 other;

Query Match 0.6%; Score 17; DB 20; Length 21170;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 ccggacgcaccatgtcg 101
|||||||||||||||||
Db 1835 CCGGACGCACCATGTCG 1819

RESULT 98
Z83005
ID X83005 standard; DNA; 29604 BP.
XX
AC X83005;
XX
DT 31-AUG-1999 (first entry)
XX
DE Partial mouse WRN genomic sequence #1.
XX
KW Mouse; WRN; Werner's syndrome; detection; diagnosis; autosomal;
KW recessive disorder; phenotype; ss.
XX
OS Mus musculus.
XX
PN WO9724435-A1.
XX
PD 10-JUL-1997.
XX
PF 30-DEC-1996; 96WO-US20785.
XX
PR 12-APR-1996; 96US-0632175.
PR 29-DEC-1995; 95US-0009409.
PR 29-DEC-1995; 95US-0580539.
PR 30-JAN-1996; 96US-0010835.
PR 30-JAN-1996; 96US-0594242.
XX
PA (DARW-) DARWIN MOLECULAR CORP.
PA (OSHI/) OSHIMA J.
XX
PI Fu Y, Mulligan J, Oshima J, Schellenberg GD, Yu C;
XX
DR WPI; 1997-363671/33.
XX
PT Isolated nucleic acid molecule encoding the WRN gene product -
PT useful for detection and treatment of Werner's syndrome, and related
PT diseases
XX
PS Claim 1; Fig 7; 153pp; English.
XX
CC This sequence represents a fragment of the genomic sequence containing
CC the coding region for the mouse WRN gene (X83004). The corresponding
CC human gene (X83001) encodes a protein related to Werner's syndrome.
CC The products can be used for the detection and treatment of Werner's
CC syndrome (WS), an autosomal recessive disorder with a complex phenotype,
CC as well as related diseases.
XX
SQ Sequence 29604 BP; 7634 A; 5861 C; 5985 G; 10123 T; 1 other;

Query Match 0.6%; Score 17; DB 18; Length 29604;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 525 aaaggaatagaactggc 541
|||||||||||||||||
Db 14826 aaaggaatagaactggc 14842

RESULT 99
Z30163/c
ID Z30163 standard; DNA; 34094 BP.
XX
AC Z30163;
XX
DT 26-JAN-2000 (first entry)
XX
DE Complete nucleotide sequence of the PAV-3 genome.
XX
KW PAV-3; defective recombinant PAV vector; live recombinant virus;
KW subunit vaccine; nucleic acid immunisation; gene therapy;
KW genetic disease; hemophilia; cystic fibrosis; cancer; viral infection;
KW acquired immune deficiency syndrome; PAV antigen; porcine pathogen; ds.
XX
OS Porcine adenovirus Type 3.
XX

PN WO9953047-A2.
XX
PD 21-OCT-1999.
XX
PF 15-APR-1999; 99WO-US08220.
XX
PR 15-APR-1998; 98US-0081882.
XX
PA (UYSA-) UNIV SASKATCHEWAN.
XX
PI Reddy PS, Tikoo SK, Babiuk LA;
XX
DR WPI; 1999-620422/53.
XX
PT New nucleic acids from the genome of porcine adenovirus-3, and derived
PT gene therapy vectors, particularly for immunization -
XX
PS Example 2; Fig 1; 87pp; English.
XX
CC The present sequence represents the complete nucleotide sequence of the
CC genome of porcine adenovirus-3 (PAV-3). The specification also describes
CC a defective recombinant PAV vector comprising inverted terminal repeats
CC (ITR), packaging sequences and at least one heterologous nucleotide
CC sequence (II), but lacking E1 functions. The defective vectors replicate
CC inefficiently in cells (other than helper cells) so are unlikely to be
CC immunogenic. Deletion of the E1 (and optionally other regions) increases
CC the size of heterologous insert that can be packaged. The PAV-3
CC polynucleotides sequences are used to produce (recombinant or defective)
CC vectors that can express heterologous proteins, e.g. for making live,
CC recombinant virus or subunit vaccines, for nucleic acid immunisation or
CC for gene therapy (e.g. of genetic diseases such as hemophilia or cystic
CC fibrosis, cancer, or viral infections, including acquired immune
CC deficiency syndrome), also for in vitro expression of recombinant
CC antigens (for antibody production), antisense RNA, ribozymes or
CC therapeutic proteins. They are also used diagnostically to detect PAV
CC antigens and/or nucleic acid. The vectors may be used in human or
CC veterinary medicine, but particularly for expressing protective
CC determinants of porcine pathogens. Regulatory regions may be used to
CC control expression of heterologous genes. Antibodies raised against PAV-3
CC polypeptides can also be used for diagnosis (to detect PAV-specific
CC antigen).
XX
SQ Sequence 34094 BP; 6240 A; 11070 C; 10693 G; 6091 T; 0 other;

Query Match 0.6%; Score 17; DB 20; Length 34094;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 330 ctcatgcaggagcacaa 346
||||||||||||||||
Db 4686 CTCATGCAGGAGCACAA 4670

RESULT 100
Z32020/c
ID Z32020 standard; DNA; 38734 BP.
XX
AC Z32020;
XX
DT 10-JAN-2000 (first entry)
XX
DE Human METH1 related EST AL021529.
XX
KW Human; METH1; METH2; anti-angiogenic; metalloprotease thrombospondin;
KW cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
KW angiogenesis inhibitor; abnormal wound healing; inflammation;
KW rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
KW diabetic retinopathy; macula degeneration; haemangioma; detection;
KW arterial-venous malformation; immune deficiency; ss.
XX
OS Homo sapiens.
XX

PN WO9937660-A1.
XX
PD 29-JUL-1999.
XX
PF 22-JAN-1999; 99WO-US01313.
XX
PR 23-JAN-1998; 98US-0072298.
PR 28-AUG-1998; 98US-0098539.
XX
PA (IRUE/) IRUELA-ARISPE L.
PA (HAST/) HASTINGS G A.
PA (RUBE/) RUBEN S M.
XX
PI Iruela-Arispe L, Hastings GA, Ruben SM;
XX
DR WPI; 1999-590684/50.
XX
PT New isolated metalloprotease thrombospondin polypeptides, useful for
PT treating hyperproliferative disorders, cancers or autoimmune disorders
PT -
XX
PS Disclosure; Page 296-321; 457pp; English.
XX
CC Z32000 and Z32001 encode, and Y49501 and Y49502 represent, human
CC metalloprotease thrombospondin (METH) proteins METH1 and METH2
CC respectively. METH1 and METH2 have been found to be potent inhibitors of
CC angiogenesis both in vitro and in vivo. They can be used for treating
CC cancer and other disorders related to angiogenesis including abnormal
CC wound healing, inflammation, rheumatoid arthritis, psoriasis,
CC endometrial bleeding disorders, diabetic retinopathy, some forms of
CC macula degeneration, haemangiomas, and arterial-venous malformations.
CC They may be useful in treating deficiencies or disorders of the immune
CC system, by activating or inhibiting the proliferation, differentiation,
CC or mobilisation (chemotaxis) of immune cells. The etiology of these
CC immune deficiencies or disorders may be genetic, somatic, such as
CC cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
CC toxins), or infectious. They can also be used to treat inflammatory
CC conditions, both chronic and acute conditions. The products can also be
CC used for detection and diagnosis. Z32002 to Z32080, and Y49503 to Y49511
CC represent sequences given in the exemplification of the present
CC invention.
XX
SQ Sequence 38734 BP; 6142 A; 13140 C; 13585 G; 5867 T; 0 other;

Query Match 0.6%; Score 17; DB 20; Length 38734;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 ggtgaccggcggctttc 27
||||||||||||||||
Db 37241 GGTGACCGGCGGCTTTC 37225

Search completed: February 18, 2001, 16:04:40
Job time: 25989 sec

OM nucleic - nucleic search, using sw model

Run on: February 18, 2001, 07:28:55 ; Search time 96.47 Seconds
(without alignments)
4941.559 Million cell updates/sec

Title: US-09-434-382-3
Perfect score: 2958
Sequence: 1 cgcgggcgtaggtgaccggc.....aataaagattgagtttgcaa 2958

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 280836 seqs, 80580151 residues

Word size : 0

Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Issued_Patents_NA:*
1: /cgn1_7/ptodata/1/ina/5A_COMB.seq:*
2: /cgn1_7/ptodata/1/ina/5B_COMB.seq:*
3: /cgn1_7/ptodata/1/ina/6_COMB.seq:*
4: /cgn1_7/ptodata/1/ina/PCTUS_COMB.seq:*
5: /cgn1_7/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	19	0.6	2955	2	US-08-867-941-9	Sequence 9, Appli
	2	19	0.6	3000	2	US-08-867-941-8	Sequence 8, Appli
	3	19	0.6	7641	2	US-08-867-941-6	Sequence 6, Appli
c	4	18	0.6	1152	3	US-09-045-186-1	Sequence 1, Appli
c	5	18	0.6	1152	3	US-09-045-186-3	Sequence 3, Appli
	6	18	0.6	2186	2	US-08-878-546-9	Sequence 9, Appli
c	7	18	0.6	2605	2	US-08-680-395-4	Sequence 4, Appli
c	8	17	0.6	289	2	US-08-967-101-23	Sequence 23, Appl
c	9	17	0.6	289	2	US-08-592-541-23	Sequence 23, Appl
c	10	17	0.6	289	3	US-09-124-698-23	Sequence 23, Appl
c	11	17	0.6	304	2	US-08-611-757-20	Sequence 20, Appl
c	12	17	0.6	304	4	PCT-US95-05980-20	Sequence 20, Appl
	13	17	0.6	1084	2	US-08-184-009-110	Sequence 110, App
	14	17	0.6	1084	2	US-08-458-356-110	Sequence 110, App
	15	17	0.6	1094	2	US-08-184-009-109	Sequence 109, App
	16	17	0.6	1094	2	US-08-458-356-109	Sequence 109, App
	17	17	0.6	1691	2	US-08-993-118-8	Sequence 8, Appli
	18	17	0.6	1691	3	US-08-845-528C-8	Sequence 8, Appli
	19	17	0.6	1816	2	US-08-951-148-2	Sequence 2, Appli
	20	17	0.6	1816	2	US-09-165-234-2	Sequence 2, Appli
	21	17	0.6	1816	3	US-09-274-570-2	Sequence 2, Appli
	22	17	0.6	2419	1	US-07-807-043B-7	Sequence 7, Appli
	23	17	0.6	2419	1	US-08-299-849B-7	Sequence 7, Appli
	24	17	0.6	2419	2	US-08-142-368A-7	Sequence 7, Appli
	25	17	0.6	2419	3	US-08-967-727-7	Sequence 7, Appli
	26	17	0.6	2420	1	US-08-465-167A-23	Sequence 23, Appl
c	27	17	0.6	3457	1	US-08-295-882-1	Sequence 1, Appli
	28	17	0.6	3796	1	US-08-343-760A-1	Sequence 1, Appli
	29	17	0.6	4488	1	US-08-441-430-1	Sequence 1, Appli
	30	17	0.6	5674	1	US-07-807-043B-8	Sequence 8, Appli
	31	17	0.6	5674	1	US-08-190-411A-1	Sequence 1, Appli
	32	17	0.6	5674	1	US-08-299-849B-8	Sequence 8, Appli
	33	17	0.6	5674	2	US-08-560-024-1	Sequence 1, Appli
	34	17	0.6	5674	2	US-08-142-368A-8	Sequence 8, Appli
	35	17	0.6	5674	3	US-08-967-727-8	Sequence 8, Appli
	36	17	0.6	6464	1	US-08-321-478-6	Sequence 6, Appli
	37	17	0.6	11236	1	US-07-853-913-1	Sequence 1, Appli
	38	17	0.6	29604	3	US-08-781-891-207	Sequence 207, App
c	39	17	0.6	68750	3	US-09-335-409-1	Sequence 1, Appli
	40	16	0.5	76	1	US-07-753-110B-12	Sequence 12, Appl
	41	16	0.5	76	1	US-08-503-730-16	Sequence 16, Appl
	42	16	0.5	76	2	US-08-507-634-13	Sequence 13, Appl
c	43	16	0.5	105	3	US-08-717-294-93	Sequence 93, Appl
	44	16	0.5	233	2	US-08-687-080-70	Sequence 70, Appl
c	45	16	0.5	252	2	US-08-630-822A-97	Sequence 97, Appl
c	46	16	0.5	252	2	US-09-005-069-97	Sequence 97, Appl
c	47	16	0.5	252	3	US-08-906-769-104	Sequence 104, App
c	48	16	0.5	252	3	US-08-906-616-104	Sequence 104, App
c	49	16	0.5	252	3	US-08-817-795-104	Sequence 104, App
c	50	16	0.5	252	3	US-08-639-075A-104	Sequence 104, App
c	51	16	0.5	252	4	PCT-US95-14442A-104	Sequence 104, App
c	52	16	0.5	294	2	US-08-611-757-98	Sequence 98, Appl
c	53	16	0.5	294	4	PCT-US95-05980-98	Sequence 98, Appl
	54	16	0.5	486	2	US-08-937-931-9	Sequence 9, Appli
	55	16	0.5	489	1	US-08-334-254-7	Sequence 7, Appli
	56	16	0.5	489	2	US-08-848-131-7	Sequence 7, Appli
	57	16	0.5	489	4	PCT-US95-14792-7	Sequence 7, Appli
	58	16	0.5	575	1	US-08-507-016-8	Sequence 8, Appli
c	59	16	0.5	578	4	PCT-US91-06418-4	Sequence 4, Appli
c	60	16	0.5	609	3	US-08-338-579A-94	Sequence 94, Appl
	61	16	0.5	654	2	US-08-911-319A-2	Sequence 2, Appli
	62	16	0.5	654	3	US-09-352-619-2	Sequence 2, Appli
	63	16	0.5	714	1	US-07-789-738-3	Sequence 3, Appli
	64	16	0.5	773	1	US-07-789-738-5	Sequence 5, Appli
	65	16	0.5	808	3	US-08-651-136C-15	Sequence 15, Appl
c	66	16	0.5	815	3	US-08-906-769-128	Sequence 128, App
c	67	16	0.5	815	3	US-08-906-616-128	Sequence 128, App
c	68	16	0.5	815	3	US-08-639-075A-128	Sequence 128, App
	69	16	0.5	1031	3	US-08-651-136C-19	Sequence 19, Appl
	70	16	0.5	1048	3	US-08-651-136C-17	Sequence 17, Appl
c	71	16	0.5	1152	2	US-08-933-750C-81	Sequence 81, Appl
c	72	16	0.5	1152	3	US-09-234-613-81	Sequence 81, Appl
	73	16	0.5	1225	2	US-08-739-485-4	Sequence 4, Appli
c	74	16	0.5	1329	3	US-08-360-758-1	Sequence 1, Appli
c	75	16	0.5	1333	3	US-08-889-425-3	Sequence 3, Appli
c	76	16	0.5	1341	3	US-09-032-372-9	Sequence 9, Appli
	77	16	0.5	1362	1	US-08-374-155A-7	Sequence 7, Appli
	78	16	0.5	1362	2	US-08-785-396-7	Sequence 7, Appli
c	79	16	0.5	1389	1	US-08-458-023B-1	Sequence 1, Appli
c	80	16	0.5	1389	3	US-09-111-556A-1	Sequence 1, Appli
c	81	16	0.5	1524	1	US-08-135-510-4	Sequence 4, Appli
c	82	16	0.5	1524	1	US-08-483-852-4	Sequence 4, Appli
c	83	16	0.5	1524	1	US-08-477-953-4	Sequence 4, Appli
	84	16	0.5	1524	1	US-08-409-122-1	Sequence 1, Appli
	85	16	0.5	1524	2	US-08-408-669-1	Sequence 1, Appli
c	86	16	0.5	1524	2	US-08-477-952-4	Sequence 4, Appli
c	87	16	0.5	1713	1	US-08-386-727-5	Sequence 5, Appli
c	88	16	0.5	1713	2	US-08-600-452A-5	Sequence 5, Appli
c	89	16	0.5	1713	2	US-08-467-948A-1	Sequence 1, Appli
c	90	16	0.5	1713	3	US-08-467-947A-1	Sequence 1, Appli
	91	16	0.5	1721	1	US-07-688-352C-13	Sequence 13, Appl
	92	16	0.5	1721	2	US-08-474-379C-13	Sequence 13, Appl
	93	16	0.5	1721	3	US-09-146-249A-13	Sequence 13, Appl
	94	16	0.5	1721	3	US-08-206-188B-13	Sequence 13, Appl
	95	16	0.5	1721	4	PCT-US91-02714-13	Sequence 13, Appl
	96	16	0.5	1766	2	US-08-481-814A-2	Sequence 2, Appli
	97	16	0.5	1790	2	US-08-993-228-1	Sequence 1, Appli
	98	16	0.5	1977	3	US-09-231-529-2	Sequence 2, Appli
	99	16	0.5	1990	3	US-09-255-911-1	Sequence 1, Appli
	100	16	0.5	2069	1	US-08-619-554-7	Sequence 7, Appli

ALIGNMENTS

```
RESULT   1
US-08-867-941-9
  Sequence 9, Application US/08867941
  Patent No. 5977337
   GENERAL INFORMATION:
     APPLICANT:  Loosmore, Sheena M
     APPLICANT:  Du, Run-Pan
     APPLICANT:  Wang, Quijun
     APPLICANT:  Yang, Yan-Ping
     APPLICANT:  Klein, Michel H
     TITLE OF INVENTION:  LACTOFERRIN RECEPTOR GENES OF MORAXELLA
     NUMBER OF SEQUENCES:  67
     CORRESPONDENCE ADDRESS:
       ADDRESSEE:  Sim & McBurney
       STREET:  6th Floor, 330 University Avenue
       CITY:  Toronto
       STATE:  Ontario
       COUNTRY:  Canada
       ZIP:  M5G 1R7
     COMPUTER READABLE FORM:
       MEDIUM TYPE:  Floppy disk
       COMPUTER:  IBM PC compatible
       OPERATING SYSTEM:  PC-DOS/MS-DOS
       SOFTWARE:  PatentIn Release #1.0, Version #1.30
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER:  US/08/867,941
       FILING DATE:  03-JUN-1997
       CLASSIFICATION:  435
     ATTORNEY/AGENT INFORMATION:
       NAME:  Stewart, Michael I
       REGISTRATION NUMBER:  24,973
       REFERENCE/DOCKET NUMBER:  1038-681 MIS:jb
     TELECOMMUNICATION INFORMATION:
       TELEPHONE:  (416) 595-1155
       TELEFAX:  (416) 595-1163
    INFORMATION FOR SEQ ID NO:  9:
     SEQUENCE CHARACTERISTICS:
       LENGTH:  2955 base pairs
       TYPE:  nucleic acid
       STRANDEDNESS:  single
       TOPOLOGY:  linear
US-08-867-941-9


  Query Match             0.6%;  Score 19;  DB 2;  Length 2955;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    811 tcttggtgctcaaagcaaa 829
          |||||||||||||||||||
Db    591 TCTTGGTGCTCAAAGCAAA 609


RESULT   2
US-08-867-941-8
  Sequence 8, Application US/08867941
  Patent No. 5977337
   GENERAL INFORMATION:
     APPLICANT:  Loosmore, Sheena M
     APPLICANT:  Du, Run-Pan
     APPLICANT:  Wang, Quijun
     APPLICANT:  Yang, Yan-Ping
     APPLICANT:  Klein, Michel H
     TITLE OF INVENTION:  LACTOFERRIN RECEPTOR GENES OF MORAXELLA
     NUMBER OF SEQUENCES:  67
     CORRESPONDENCE ADDRESS:
       ADDRESSEE:  Sim & McBurney
       STREET:  6th Floor, 330 University Avenue
       CITY:  Toronto
       STATE:  Ontario
       COUNTRY:  Canada
       ZIP:  M5G 1R7
     COMPUTER READABLE FORM:
       MEDIUM TYPE:  Floppy disk
       COMPUTER:  IBM PC compatible
       OPERATING SYSTEM:  PC-DOS/MS-DOS
       SOFTWARE:  PatentIn Release #1.0, Version #1.30
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER:  US/08/867,941
       FILING DATE:  03-JUN-1997
       CLASSIFICATION:  435
     ATTORNEY/AGENT INFORMATION:
       NAME:  Stewart, Michael I
       REGISTRATION NUMBER:  24,973
       REFERENCE/DOCKET NUMBER:  1038-681 MIS:jb
     TELECOMMUNICATION INFORMATION:
       TELEPHONE:  (416) 595-1155
       TELEFAX:  (416) 595-1163
    INFORMATION FOR SEQ ID NO:  8:
     SEQUENCE CHARACTERISTICS:
       LENGTH:  3000 base pairs
       TYPE:  nucleic acid
       STRANDEDNESS:  single
       TOPOLOGY:  linear
US-08-867-941-8


  Query Match             0.6%;  Score 19;  DB 2;  Length 3000;
  Best Local Similarity   100.0%;  Pred. No. 7.5;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    811 tcttggtgctcaaagcaaa 829
          |||||||||||||||||||
Db    636 TCTTGGTGCTCAAAGCAAA 654


RESULT   3
US-08-867-941-6
  Sequence 6, Application US/08867941
  Patent No. 5977337
   GENERAL INFORMATION:
     APPLICANT:  Loosmore, Sheena M
     APPLICANT:  Du, Run-Pan
     APPLICANT:  Wang, Quijun
     APPLICANT:  Yang, Yan-Ping
     APPLICANT:  Klein, Michel H
     TITLE OF INVENTION:  LACTOFERRIN RECEPTOR GENES OF MORAXELLA
     NUMBER OF SEQUENCES:  67
     CORRESPONDENCE ADDRESS:
       ADDRESSEE:  Sim & McBurney
       STREET:  6th Floor, 330 University Avenue
       CITY:  Toronto
       STATE:  Ontario
       COUNTRY:  Canada
       ZIP:  M5G 1R7
     COMPUTER READABLE FORM:
       MEDIUM TYPE:  Floppy disk
       COMPUTER:  IBM PC compatible
       OPERATING SYSTEM:  PC-DOS/MS-DOS
       SOFTWARE:  PatentIn Release #1.0, Version #1.30
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER:  US/08/867,941
       FILING DATE:  03-JUN-1997
       CLASSIFICATION:  435
     ATTORNEY/AGENT INFORMATION:
       NAME:  Stewart, Michael I
       REGISTRATION NUMBER:  24,973
       REFERENCE/DOCKET NUMBER:  1038-681 MIS:jb
     TELECOMMUNICATION INFORMATION:
       TELEPHONE:  (416) 595-1155
       TELEFAX:  (416) 595-1163
    INFORMATION FOR SEQ ID NO:  6:
```

```
    SEQUENCE CHARACTERISTICS:
      LENGTH:  7641 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
S-08-867-941-6


  Query Match             0.6%;  Score 19;  DB 2;  Length 7641;
  Best Local Similarity   100.0%;  Pred. No. 7.6;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    811 tcttggtgctcaaagcaaa 829
          ||||||||||||||||||
Db   3656 TCTTGGTGCTCAAAGCAAA 3674


ESULT   4
S-09-045-186-1/c
 Sequence 1, Application US/09045186
 Patent No. 6087154
  GENERAL INFORMATION:
    APPLICANT:  Baez, Melvyn
    APPLICANT:  Cates, Carolyn A.
    TITLE OF INVENTION:  RHESUS NEUROPEPTIDE Y1 RECEPTOR
    NUMBER OF SEQUENCES:  3
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Eli Lilly and Company
      STREET:  Lill Corporate Center
      CITY:  Indianapolis
      STATE:  Indiana
      COUNTRY:  United States of America
      ZIP:  46285
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/09/045,186
      FILING DATE:
      CLASSIFICATION:
    ATTORNEY/AGENT INFORMATION:
      NAME:  Gaylo, Paul J.
      REGISTRATION NUMBER:  36,808
      REFERENCE/DOCKET NUMBER:  P-11376
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (317) 276-0756
      TELEFAX:  (317) 276-3861
  INFORMATION FOR SEQ ID NO:  1:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  1152 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    MOLECULE TYPE:  cDNA
    FEATURE:
      NAME/KEY:  CDS
      LOCATION:  1..1152
S-09-045-186-1


  Query Match             0.6%;  Score 18;  DB 3;  Length 1152;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2802 aaagaagtcttggaaaca 2819
          ||||||||||||||||||
Db   1091 AAAGAAGTCTTGGAAACA 1074


ESULT   5
S-09-045-186-3/c
; Sequence 3, Application US/09045186
; Patent No. 6087154
; GENERAL INFORMATION:
;   APPLICANT:  Baez, Melvyn
;   APPLICANT:  Cates, Carolyn A.
;   TITLE OF INVENTION:  RHESUS NEUROPEPTIDE Y1 RECEPTOR
;   NUMBER OF SEQUENCES:  3
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Eli Lilly and Company
;     STREET:  Lill Corporate Center
;     CITY:  Indianapolis
;     STATE:  Indiana
;     COUNTRY:  United States of America
;     ZIP:  46285
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/09/045,186
;     FILING DATE:
;     CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Gaylo, Paul J.
;     REGISTRATION NUMBER:  36,808
;     REFERENCE/DOCKET NUMBER:  P-11376
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (317) 276-0756
;     TELEFAX:  (317) 276-3861
; INFORMATION FOR SEQ ID NO:  3:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  1152 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  mRNA
US-09-045-186-3


  Query Match             0.6%;  Score 18;  DB 3;  Length 1152;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2802 aaagaagtcttggaaaca 2819
          ||||||||||||||||||
Db   1091 AAAGAAGTCTTGGAAACA 1074


RESULT   6
US-08-878-546-9
; Sequence 9, Application US/08878546
; Patent No. 5952463
; GENERAL INFORMATION:
;   APPLICANT:  SHIBANO, YUJI
;   APPLICANT:  KIKUCHI, NORIHISA
;   APPLICANT:  ODA, KOHEI
;   TITLE OF INVENTION:  NOVEL PROTEINASE INHIBITOR AND
;   TITLE OF INVENTION:  GENE ENCODING THE INHIBITOR
;   NUMBER OF SEQUENCES:  10
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  STEINBERG, RASKIN & DAVIDSON P.C.
;     STREET:  1140 AVENUE OF THE AMERICAS
;     CITY:  NEW YORK
;     STATE:  NY
;     COUNTRY:  USA
;     ZIP:  10036
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
```

```
      APPLICATION NUMBER:  US/08/878,546
      FILING DATE:  19-JUN-1997
      CLASSIFICATION:  435
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  JP 158677/1996
      FILING DATE:  19-JUN-1996
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  JP 224104/1996
      FILING DATE:  26-AUG-1996
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  JP 48101/1997
      FILING DATE:  03-MAR-1997
    ATTORNEY/AGENT INFORMATION:
      NAME:  DAVIDSON, CLIFFORD M.
      REGISTRATION NUMBER:  32,728
      REFERENCE/DOCKET NUMBER:  382.1009
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (212)-768-3800
      TELEFAX:  (212)382-2124
  INFORMATION FOR SEQ ID NO:  9:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  2186 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  double
      TOPOLOGY:  linear
    MOLECULE TYPE:  DNA (genomic)
    HYPOTHETICAL:  NO
    ANTI-SENSE:  NO
    ORIGINAL SOURCE:
      ORGANISM:  STREPTOMYCES PLATENSIS
      STRAIN:  Q268
    FEATURE:
      NAME/KEY:  CDS
      LOCATION:  1477..1911
S-08-878-546-9


  Query Match              0.6%;  Score 18;  DB 2;  Length 2186;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

y   1652 ggtcctgggcaccctggc 1669
         ||||||||||||||||||
b   1518 GGTCCTGGGCACCCTGGC 1535


ESULT   7
S-08-680-395-4/c
  Sequence 4, Application US/08680395
  Patent No. 5892010
   GENERAL INFORMATION:
    APPLICANT:  Gray, Joe W.
    APPLICANT:  Collins, Colin
    APPLICANT:  Hwang, Soo-in
    APPLICANT:  Godfrey, Tony
    APPLICANT:  Kowbel, David
    APPLICANT:  Rommens, Johanna
    TITLE OF INVENTION:  Genes from the 20q13 Amplicon and Their
    TITLE OF INVENTION:  Uses
    NUMBER OF SEQUENCES:  40
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Townsend and Townsend and Crew LLP
      STREET:  Two Embarcadero Center, Eighth Floor
      CITY:  San Francisco
      STATE:  California
      COUNTRY:  USA
      ZIP:  94111-3834
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/680,395
;     FILING DATE:  15-JUL-1996
;     CLASSIFICATION:  435
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  Bastian, Kevin L.
;     REGISTRATION NUMBER:  34,774
;     REFERENCE/DOCKET NUMBER:  02307O-068900US
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (415) 576-0200
;     TELEFAX:  (415) 576-0300
; INFORMATION FOR SEQ ID NO:  4:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  2605 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  cDNA
;   FEATURE:
;     NAME/KEY:  -
;     LOCATION:  1..2605
;     OTHER INFORMATION:  /note= "cDNA clone cc43 of 4 kb
;     OTHER INFORMATION:  transcript"
US-08-680-395-4


  Query Match              0.6%;  Score 18;  DB 2;  Length 2605;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    448 ttccaaagtgtgtacttt 465
          ||||||||||||||||||
Db   1421 TTCCAAAGTGTGTACTTT 1404


RESULT   8
US-08-967-101-23/c
; Sequence 23, Application US/08967101
; Patent No. 5840540
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/967,101
;      FILING DATE:  10-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/592,541
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  289 base pairs
;      TYPE:  nucleic acid
```

```
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    MOLECULE TYPE:  DNA (genomic)
US-08-967-101-23


  Query Match             0.6%;  Score 17;  DB 2;  Length 289;
  Best Local Similarity   100.0%;  Pred. No. 71;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2131 aagaggaagcagtggaa 2147
          |||||||||||||||||
Db    267 AAGAGGAAGCAGTGGAA 251


RESULT   9
US-08-592-541-23/c
; Sequence 23, Application US/08592541
; Patent No. 5986054
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,541
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  289 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-592-541-23


  Query Match             0.6%;  Score 17;  DB 2;  Length 289;
  Best Local Similarity   100.0%;  Pred. No. 71;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2131 aagaggaagcagtggaa 2147
          |||||||||||||||||
Db    267 AAGAGGAAGCAGTGGAA 251


RESULT  10
US-09-124-698-23/c
; Sequence 23, Application US/09124698
; Patent No. 6117978
;  GENERAL INFORMATION:
;    APPLICANT:  ST. GEORGE-HYSLOP, PETER H
;    APPLICANT:  ROMMENS, JOHANNA M
;    APPLICANT:  FRASER, PAUL E
;    TITLE OF INVENTION:  GENETIC SEQUENCES AND PROTEINS RELATED
;    TITLE OF INVENTION:  TO ALZHEIMER'S DISEASE
;    NUMBER OF SEQUENCES:  183
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TESTA, HURWITZ & THIBEAULT
;      STREET:  High Street Tower - 125 High Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/124,698
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/592,541
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pitcher, Edmund R.
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 248-7000
;      TELEFAX:  (617) 248-7100
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  289 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-124-698-23


  Query Match             0.6%;  Score 17;  DB 3;  Length 289;
  Best Local Similarity   100.0%;  Pred. No. 71;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2131 aagaggaagcagtggaa 2147
          |||||||||||||||||
Db    267 AAGAGGAAGCAGTGGAA 251


RESULT  11
US-08-611-757-20/c
; Sequence 20, Application US/08611757
; Patent No. 5859230
;  GENERAL INFORMATION:
;    APPLICANT:  Kim, Jungsuh P.
;    APPLICANT:  Reyes, Gregory R.
;    APPLICANT:  Wages, John
;    APPLICANT:  Zhang-Keck, Zhen-Yang
;    APPLICANT:  Young, LaVonne
;    TITLE OF INVENTION:  No. 5859230-A/No. 5859230-B/No. 5859230-C/No. 5859230-D/No.
;    TITLE OF INVENTION:  Agents and Molecular Cloning Thereof
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  350 Cambridge Avenue, Suite 250
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
```

```
      APPLICATION NUMBER:  US/08/611,757
      FILING DATE:
      CLASSIFICATION:  435
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  08/246,985
      FILING DATE:  20-MAY-1994
      APPLICATION NUMBER:  US 025,396
      FILING DATE:  24-FEB-1993
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  US 07/922,493
      FILING DATE:  30-JUL-1992
    ATTORNEY/AGENT INFORMATION:
      NAME:  Fabian, Gary R.
      REGISTRATION NUMBER:  33,875
      REFERENCE/DOCKET NUMBER:  4600-0201
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (415) 324-0880
      TELEFAX:  (415) 324-0960
  INFORMATION FOR SEQ ID NO:  20:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  304 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  double
      TOPOLOGY:  linear
    MOLECULE TYPE:  DNA
    HYPOTHETICAL:  NO
    ANTI-SENSE:  NO
    ORIGINAL SOURCE:
      INDIVIDUAL ISOLATE:  My 190 Clone D30
S-08-611-757-20


  Query Match             0.6%;  Score 17;  DB 2;  Length 304;
  Best Local Similarity   100.0%;  Pred. No. 71;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     132 cgccgcgagcggccgcg 148
           |||||||||||||||||
Db      22 CGCCGCGAGCGGCCGCG 6


RESULT  12
PCT-US95-05980-20/c
  Sequence 20, Application PC/TUS9505980
  GENERAL INFORMATION:
    APPLICANT:
    APPLICANT:
    TITLE OF INVENTION:  Non-A/Non-B/Non-C/Non-D/Non-E Hepatitis
    TITLE OF INVENTION:  Agents and Molecular Cloning Thereof
    NUMBER OF SEQUENCES:  106
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Dehlinger & Associates
      STREET:  350 Cambridge Avenue, Suite 250
      CITY:  Palo Alto
      STATE:  CA
      COUNTRY:  USA
      ZIP:  94306
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  PCT/US95/05980
      FILING DATE:
      CLASSIFICATION:
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  US 08/246,986
      FILING DATE:  20-MAY-1994
    ATTORNEY/AGENT INFORMATION:
      NAME:  Fabian, Gary R.
      REGISTRATION NUMBER:  33,875
      REFERENCE/DOCKET NUMBER:  4600-0201.49
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (415) 324-0880
      TELEFAX:  (415) 324-0960
  INFORMATION FOR SEQ ID NO:  20:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  304 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  double
      TOPOLOGY:  linear
    MOLECULE TYPE:  DNA
    HYPOTHETICAL:  NO
    ANTI-SENSE:  NO
    ORIGINAL SOURCE:
      INDIVIDUAL ISOLATE:  My 190 Clone D30
PCT-US95-05980-20


  Query Match             0.6%;  Score 17;  DB 4;  Length 304;
  Best Local Similarity   100.0%;  Pred. No. 71;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     132 cgccgcgagcggccgcg 148
           |||||||||||||||||
Db      22 CGCCGCGAGCGGCCGCG 6


RESULT  13
US-08-184-009-110
; Sequence 110, Application US/08184009
; Patent No. 5833975
;  GENERAL INFORMATION:
;    APPLICANT:  Paoletti, Enzo
;    APPLICANT:  Tartaglia, James
;    APPLICANT:  Cox, William I.
;    TITLE OF INVENTION:  RECOMBINANT VIRUS IMMUNOTHERAPY
;    NUMBER OF SEQUENCES:  217
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/184,009
;      FILING DATE:  19-JAN-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Frommer, William S.
;      REGISTRATION NUMBER:  25,506
;      REFERENCE/DOCKET NUMBER:  454310-2530
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;      TELEX:  425066CURTMS
;  INFORMATION FOR SEQ ID NO:  110:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1084 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-184-009-110


  Query Match             0.6%;  Score 17;  DB 2;  Length 1084;
  Best Local Similarity   100.0%;  Pred. No. 72;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy   1652 ggtcctgggcaccctgg 1668
          |||||||||||||||||
Db    292 GGTCCTGGGCACCCTGG 308


RESULT  14
US-08-458-356-110
  Sequence 110, Application US/08458356
  Patent No. 5942235
   GENERAL INFORMATION:
     APPLICANT:  Paoletti, Enzo
     APPLICANT:  Tartaglia, James
     APPLICANT:  Cox, William I.
     TITLE OF INVENTION:  RECOMBINANT VIRUS IMMUNOTHERAPY
     NUMBER OF SEQUENCES:  217
     CORRESPONDENCE ADDRESS:
       ADDRESSEE:  Curtis, Morris & Safford
       STREET:  530 Fifth Avenue
       CITY:  New York
       STATE:  NY
       COUNTRY:  USA
       ZIP:  10036
     COMPUTER READABLE FORM:
       MEDIUM TYPE:  Floppy disk
       COMPUTER:  IBM PC compatible
       OPERATING SYSTEM:  PC-DOS/MS-DOS
       SOFTWARE:  PatentIn Release #1.0, Version #1.25
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER:  US/08/458,356
       FILING DATE:  02-JUN-1995
       CLASSIFICATION:  424
     PRIOR APPLICATION DATA:
       APPLICATION NUMBER:  US 08/184,009
       FILING DATE:  19-JAN-1994
     ATTORNEY/AGENT INFORMATION:
       NAME:  Frommer, William S.
       REGISTRATION NUMBER:  25,506
       REFERENCE/DOCKET NUMBER:  454310-2530
     TELECOMMUNICATION INFORMATION:
       TELEPHONE:  (212) 840-3333
       TELEFAX:  (212) 840-0712
       TELEX:  425066CURTMS
   INFORMATION FOR SEQ ID NO:  110:
     SEQUENCE CHARACTERISTICS:
       LENGTH:  1084 base pairs
       TYPE:  nucleic acid
       STRANDEDNESS:  single
       TOPOLOGY:  linear
     MOLECULE TYPE:  cDNA
US-08-458-356-110


  Query Match             0.6%;  Score 17;  DB 2;  Length 1084;
  Best Local Similarity   100.0%;  Pred. No. 72;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1652 ggtcctgggcaccctgg 1668
          |||||||||||||||||
Db    292 GGTCCTGGGCACCCTGG 308


RESULT  15
US-08-184-009-109
  Sequence 109, Application US/08184009
  Patent No. 5833975
   GENERAL INFORMATION:
     APPLICANT:  Paoletti, Enzo
     APPLICANT:  Tartaglia, James
     APPLICANT:  Cox, William I.
     TITLE OF INVENTION:  RECOMBINANT VIRUS IMMUNOTHERAPY
     NUMBER OF SEQUENCES:  217
     CORRESPONDENCE ADDRESS:
       ADDRESSEE:  Curtis, Morris & Safford
       STREET:  530 Fifth Avenue
       CITY:  New York
       STATE:  NY
       COUNTRY:  USA
       ZIP:  10036
     COMPUTER READABLE FORM:
       MEDIUM TYPE:  Floppy disk
       COMPUTER:  IBM PC compatible
       OPERATING SYSTEM:  PC-DOS/MS-DOS
       SOFTWARE:  PatentIn Release #1.0, Version #1.25
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER:  US/08/184,009
       FILING DATE:  19-JAN-1994
       CLASSIFICATION:  435
     ATTORNEY/AGENT INFORMATION:
       NAME:  Frommer, William S.
       REGISTRATION NUMBER:  25,506
       REFERENCE/DOCKET NUMBER:  454310-2530
     TELECOMMUNICATION INFORMATION:
       TELEPHONE:  (212) 840-3333
       TELEFAX:  (212) 840-0712
       TELEX:  425066CURTMS
   INFORMATION FOR SEQ ID NO:  109:
     SEQUENCE CHARACTERISTICS:
       LENGTH:  1094 base pairs
       TYPE:  nucleic acid
       STRANDEDNESS:  single
       TOPOLOGY:  linear
     MOLECULE TYPE:  cDNA
US-08-184-009-109


  Query Match             0.6%;  Score 17;  DB 2;  Length 1094;
  Best Local Similarity   100.0%;  Pred. No. 72;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1652 ggtcctgggcaccctgg 1668
          |||||||||||||||||
Db    314 GGTCCTGGGCACCCTGG 330


RESULT  16
US-08-458-356-109
  Sequence 109, Application US/08458356
  Patent No. 5942235
   GENERAL INFORMATION:
     APPLICANT:  Paoletti, Enzo
     APPLICANT:  Tartaglia, James
     APPLICANT:  Cox, William I.
     TITLE OF INVENTION:  RECOMBINANT VIRUS IMMUNOTHERAPY
     NUMBER OF SEQUENCES:  217
     CORRESPONDENCE ADDRESS:
       ADDRESSEE:  Curtis, Morris & Safford
       STREET:  530 Fifth Avenue
       CITY:  New York
       STATE:  NY
       COUNTRY:  USA
       ZIP:  10036
     COMPUTER READABLE FORM:
       MEDIUM TYPE:  Floppy disk
       COMPUTER:  IBM PC compatible
       OPERATING SYSTEM:  PC-DOS/MS-DOS
       SOFTWARE:  PatentIn Release #1.0, Version #1.25
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER:  US/08/458,356
       FILING DATE:  02-JUN-1995
       CLASSIFICATION:  424
     PRIOR APPLICATION DATA:
       APPLICATION NUMBER:  US 08/184,009
       FILING DATE:  19-JAN-1994
     ATTORNEY/AGENT INFORMATION:
       NAME:  Frommer, William S.
```

```
      REGISTRATION NUMBER:  25,506
      REFERENCE/DOCKET NUMBER:  454310-2530
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (212) 840-3333
      TELEFAX:  (212) 840-0712
      TELEX:  425066CURTMS
  INFORMATION FOR SEQ ID NO:  109:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  1094 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    MOLECULE TYPE:  cDNA
US-08-458-356-109


  Query Match             0.6%;  Score 17;  DB 2;  Length 1094;
  Best Local Similarity   100.0%;  Pred. No. 72;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1652 ggtcctgggcaccctgg 1668
          |||||||||||||||||
Db    314 GGTCCTGGGCACCCTGG 330


RESULT  17
US-08-993-118-8
; Sequence 8, Application US/08993118
; Patent No. 5997872
;  GENERAL INFORMATION:
;    APPLICANT:  LUCAS, Sophie;
;    APPLICANT:  DE SMET, Charles;
;    APPLICANT:  BOON-FALLEUR, Thierry
;    TITLE OF INVENTION:  ISOLATED NUCLEIC ACID MOLECULE CODING FOR TUMOR
;    TITLE OF INVENTION:  REJECTION ANTIGEN PRECURSOR MAGE-C1 AND USES
;    TITLE OF INVENTION:  THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 360 kb storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/993,118
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/845,528
;      FILING DATE:  April 25, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mary Anne Schofield
;      REGISTRATION NUMBER:  36,669
;      REFERENCE/DOCKET NUMBER:  LUD 5455
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1691 base pairs
;      TYPE:  nucleotides
;      STRANDEDNESS:  single stranded
;      TOPOLOGY:  linear
US-08-993-118-8


  Query Match             0.6%;  Score 17;  DB 2;  Length 1691;
  Best Local Similarity   100.0%;  Pred. No. 72;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1652 ggtcctgggcaccctgg 1668
          |||||||||||||||||
Db    317 GGTCCTGGGCACCCTGG 333


RESULT  18
US-08-845-528C-8
; Sequence 8, Application US/08845528C
; Patent No. 6027924
;  GENERAL INFORMATION:
;    APPLICANT:  LUCAS, Sophie;
;    APPLICANT:  DE SMET, Charles;
;    APPLICANT:  BOON-FALLEUR, Thierry
;    TITLE OF INVENTION:  ISOLATED NUCLEIC ACID MOLECULE CODING FOR TUMOR
;    TITLE OF INVENTION:  REJECTION ANTIGEN PRECURSOR MAGE-C1 AND USES
;    TITLE OF INVENTION:  THEREOF
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 360 kb storage
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/845,528C
;      FILING DATE:  April 25, 1997
;      CLASSIFICATION:  4335
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mary Anne Schofield
;      REGISTRATION NUMBER:  36,669
;      REFERENCE/DOCKET NUMBER:  LUD 5455
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1691 base pairs
;      TYPE:  nucleotides
;      STRANDEDNESS:  single stranded
;      TOPOLOGY:  linear
US-08-845-528C-8


  Query Match             0.6%;  Score 17;  DB 3;  Length 1691;
  Best Local Similarity   100.0%;  Pred. No. 72;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1652 ggtcctgggcaccctgg 1668
          |||||||||||||||||
Db    317 GGTCCTGGGCACCCTGG 333


RESULT  19
US-08-951-148-2
; Sequence 2, Application US/08951148
; Patent No. 5871973
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  CELL DIVISION REGULATORS
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
```

```
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/951,148
;      FILING DATE:  Herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0407 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1816 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  SPLNFZT01
;      CLONE:  26459
US-08-951-148-2


  Query Match             0.6%;  Score 17;  DB 2;  Length 1816;
  Best Local Similarity   100.0%;  Pred. No. 72;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     804 ggaaacttcttggtgct 820
           |||||||||||||||||
Db     717 GGAAACTTCTTGGTGCT 733


RESULT  20
US-09-165-234-2
; Sequence 2, Application US/09165234
; Patent No. 5928899
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  CELL DIVISION REGULATORS
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/165,234
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/951,148
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0407 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1816 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  SPLNFZT01
;      CLONE:  26459
US-09-165-234-2


  Query Match             0.6%;  Score 17;  DB 2;  Length 1816;
  Best Local Similarity   100.0%;  Pred. No. 72;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     804 ggaaacttcttggtgct 820
           |||||||||||||||||
Db     717 GGAAACTTCTTGGTGCT 733


RESULT  21
US-09-274-570-2
; Sequence 2, Application US/09274570
; Patent No. 6121019
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  CELL DIVISION REGULATORS
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/274,570
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/951,148
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0407 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
```

LENGTH: 1816 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SPLNFZT01
CLONE: 26459
S-09-274-570-2

Query Match 0.6%; Score 17; DB 3; Length 1816;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 804 ggaaacttcttggtgct 820
||||||||||||||||
Db 717 GGAAACTTCTTGGTGCT 733

RESULT 22
S-07-807-043B-7
Sequence 7, Application US/07807043B
Patent No. 5342774
GENERAL INFORMATION:
APPLICANT: Boon, Thierry, Van den Eynde, Beno t
TITLE OF INVENTION: Tumor Rejection Antigen Precursors, Tumor
TITLE OF INVENTION: Rejection Antigens and Uses Thereof
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/807,043B
FILING DATE: 19911212
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/764,364
FILING DATE: 23-SEPTEMBER-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/728,838
FILING DATE: 9-JULY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/705,702
FILING DATE: 23-May-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5342774man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 253.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2419 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: singular
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
S-07-807-043B-7

Query Match 0.6%; Score 17; DB 1; Length 2419;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1652 ggtcctgggcaccctgg 1668
||||||||||||||||
Db 739 GGTCCTGGGCACCCTGG 755

RESULT 23
US-08-299-849B-7
; Sequence 7, Application US/08299849B
; Patent No. 5612201
; GENERAL INFORMATION:
; APPLICANT: De Plaen, Etienne; Boon-Falleur, Thierry;
; APPLICANT: Leth , Bernard; Szikora, Jean-Pierre; De Smet, Charles;
; APPLICANT: Chomez, Patrick
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Useful In
; TITLE OF INVENTION: Determining Expression Of A Tumor Antigen Precursor
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/299,849B
; FILING DATE: 1-SEPTEMBER-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/037,230
; FILING DATE: 26-MARCH-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/04354
; FILING DATE: 22-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/807,043
; FILING DATE: 12-DECEMBER-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/764,364
; FILING DATE: 23-SEPTEMBER-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/728,838
; APPLICATION NUMBER: 9-JULY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/705,702
; FILING DATE: 23-May-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5612201man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5355
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2419 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
US-08-299-849B-7

Query Match 0.6%; Score 17; DB 1; Length 2419;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1652 ggtcctgggcaccctgg 1668
||||||||||||||||
Db 739 GGTCCTGGGCACCCTGG 755

```
ESULT  24
S-08-142-368A-7
 Sequence 7, Application US/08142368A
 Patent No. 5925729
  GENERAL INFORMATION:
    APPLICANT:  Boon-Falleur, Thierry; Van der Bruggen, Thierry;
    APPLICANT:  Van den Eynde, Beno t; Van Pel, Aline; De Plaen, Etienne;
    APPLICANT:  Lurquin, Christophe; Chomez, Patrick; Traversari, Catia
    TITLE OF INVENTION:  Tumor Rejection Antigen Precursors, Tumor
    TITLE OF INVENTION:  Rejection Antigens and Uses Thereof
    NUMBER OF SEQUENCES:  26
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Felfe & Lynch
      STREET:  805 Third Avenue
      CITY:  New York City
      STATE:  New York
      ZIP:  10022
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
      COMPUTER:  IBM
      OPERATING SYSTEM:  PC-DOS
      SOFTWARE:  Wordperfect
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/142,368A
      FILING DATE:  02-MAY-1994
      CLASSIFICATION:  435
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  PCT/US92/04354
      FILING DATE:  22-MAY-1992
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  07/807,043
      FILING DATE:  12-DECEMBER-1991
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  07/764,364
      FILING DATE:  23-SEPTEMBER-1991
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  07/728,838
      APPLICATION NUMBER:  9-JULY-1991
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  07/705,702
      FILING DATE:  23-May-1991
    ATTORNEY/AGENT INFORMATION:
      NAME:  Hanson, No. 5925729man D.
      REGISTRATION NUMBER:  30,946
      REFERENCE/DOCKET NUMBER:  LUD 5253.4-US
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (212) 688-9200
      TELEFAX:  (212) 838-3884
  INFORMATION FOR SEQ ID NO:  7:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  2419 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    MOLECULE TYPE:  genomic DNA
S-08-142-368A-7


  Query Match             0.6%;  Score 17;  DB 2;  Length 2419;
  Best Local Similarity   100.0%;  Pred. No. 73;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1652 ggtcctgggcaccctgg 1668
          |||||||||||||||||
Db    739 GGTCCTGGGCACCCTGG 755


ESULT  25
S-08-967-727-7
 Sequence 7, Application US/08967727
 Patent No. 6025474
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Gaugler, B atrice; Van den Eynde, Beno t;
;    APPLICANT:  van der Bruggen, Pierre;  Boon-Falleur, Thierry
;    TITLE OF INVENTION:  Isolated Nucleic Acid Molecules Coding For
;    TITLE OF INVENTION:  Tumor Rejection Antigen Precursor Mage-3 And Uses Thereof
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/967,727
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/037,230
;      FILING DATE:  26-MARCH-1993
;      APPLICATION NUMBER:  PCT/US92/04354
;      FILING DATE:  22-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/807,043
;      FILING DATE:  12-DECEMBER-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/764,365
;      FILING DATE:  23-SEPTEMBER-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/728,838
;      FILING DATE:  9-JULY-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/705,702
;      FILING DATE:  23-MAY-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 6025474man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  LUD 5353
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2419 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  genomic DNA
US-08-967-727-7


  Query Match             0.6%;  Score 17;  DB 3;  Length 2419;
  Best Local Similarity   100.0%;  Pred. No. 73;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1652 ggtcctgggcaccctgg 1668
          |||||||||||||||||
Db    739 GGTCCTGGGCACCCTGG 755


RESULT  26
US-08-465-167A-23-
; Sequence 23, Application US/08465167A
; Patent No. 5750395
;  GENERAL INFORMATION:
;    APPLICANT:  Fikes, John D.
;    APPLICANT:  Livingston, Brian D.
;    APPLICANT:  Sette, Alessandro D.
;    APPLICANT:  Sidney, John C.
```

```
    TITLE OF INVENTION:  DNA ENCODING MAGE-1 C-TERMINAL
    TITLE OF INVENTION:  IMMUNOGENIC PEPTIDES (as amended)
    NUMBER OF SEQUENCES:  51
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Townsend and Townsend and Crew LLP
      STREET:  Two Embarcadero Center, 8th Floor
      CITY:  San Francisco
      STATE:  CA
      COUNTRY:  USA
      ZIP:  98111
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/465,167A
      FILING DATE:  05-JUN-1995
      CLASSIFICATION:  435
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  US 08/103,623
      FILING DATE:  06-AUG-1993
    ATTORNEY/AGENT INFORMATION:
      NAME:  Parmelee, Steven W.
      REGISTRATION NUMBER:  31,990
      REFERENCE/DOCKET NUMBER:  14137-60-1
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  206-467-9600
      TELEFAX:  415-576-0300
  INFORMATION FOR SEQ ID NO:  23:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  2420 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    MOLECULE TYPE:  protein
    FEATURE:
      NAME/KEY:  CDS
      LOCATION:  626..1552
S-08-465-167A-23


  Query Match              0.6%;  Score 17;  DB 1;  Length 2420;
  Best Local Similarity   100.0%;  Pred. No. 73;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

y   1652 ggtcctgggcaccctgg 1668
          |||||||||||||||||
    739 GGTCCTGGGCACCCTGG 755


ESULT  27
S-08-295-882-1/c
 Sequence 1, Application US/08295882
 Patent No. 5569833
  GENERAL INFORMATION:
    APPLICANT:
    TITLE OF INVENTION:  METHOD FOR ENHANCING
    TITLE OF INVENTION:  PLANT PRECOCITY AND/OR REDUCING THE
    TITLE OF INVENTION:  STORED NITRATE CONTENT OF A PLANT
    NUMBER OF SEQUENCES:  1
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Christie, Parker & Hale
      STREET:  P.O. Box 7068
      CITY:  Pasadena
      STATE:  CA
      COUNTRY:  USA
      ZIP:  91109-7068
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  WORD PERFECT 5.1 release
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/295,882
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR 93/00222
;      FILING DATE:  March 5, 1993
;      APPLICATION NUMBER:  92 02658
;      FILING DATE:  March 5, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Prout, D. Bruce
;      REGISTRATION NUMBER:  20,958
;      REFERENCE/DOCKET NUMBER:  27209/DBP
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3457 base pairs
;      TYPE:  nucleotide with corresponding
;      TYPE:  protein
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    ORIGINAL SOURCE:
;    ORIGINAL SOURCE:  (1.vi.A)  ORGANISM:      Nicotiana tabacum
;    ORIGINAL SOURCE:  (1.vi.B)  CELL LINE:     N. tabacum cv. Xanthi
;    ORIGINAL SOURCE:  XHFD 8
;    IMMEDIATE SOURCE:  leaf
;    FEATURE:
;      NAME/KEY:  Nitrate reductase
;      LOCATION:  from 1 to 143 bp: Leader
;      LOCATION:  non translated 5 sequence (leader)
;      LOCATION:  from 144 to 2855 bp: coding sequence
;      LOCATION:  for nitrate reductase apoenzyme
;      LOCATION:  from 2856 to 3457 bp: non translated
;      LOCATION:  3  sequence
US-08-295-882-1


  Query Match              0.6%;  Score 17;  DB 1;  Length 3457;
  Best Local Similarity   100.0%;  Pred. No. 73;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    909 gaaggaagagagatttt 925
          |||||||||||||||||
Db     94 GAAGGAAGAGAGATTTT 78


RESULT  28
US-08-343-760A-1
; Sequence 1, Application US/08343760A
; Patent No. 5679783
;  GENERAL INFORMATION:
;    APPLICANT:  De Robertis, Edward M
;    APPLICANT:  Sasai, Yoshiki
;    TITLE OF INVENTION:  Tissue Differentiation Affecting
;    TITLE OF INVENTION:  Factor and Composition
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Majestic, Parsons, Siebert & Hsue
;      STREET:  Four Embarcadero Center, Suite 1450
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/343,760A
;      FILING DATE:  22-NOV-1994
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Siebert, J. Suzanne
```

```
;      REGISTRATION NUMBER:  28,758
;      REFERENCE/DOCKET NUMBER:  3100.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 363-5556
;      TELEFAX:  (415) 362-5418
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3796 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;S-08-343-760A-1


  Query Match             0.6%;  Score 17;  DB 1;  Length 3796;
  Best Local Similarity   100.0%;  Pred. No. 73;
  Matches   17;  Conservative    0;  Mismatches    0;  Indexels    0;  Gaps    0;

;y    270 ctctacgtcttctccga 286
          |||||||||||||||||
;b   1036 CTCTACGTCTTCTCCGA 1052


;ESULT  29
;S-08-441-430-1
;  Sequence 1, Application US/08441430
;  Patent No. 5681942
;  GENERAL INFORMATION:
;    APPLICANT:  Buchwald, Manuel
;    APPLICANT:  Strathdee, Craig A.
;    APPLICANT:  Wevrick, Rachel
;    APPLICANT:  Mathew, Christopher George Porter
;    TITLE OF INVENTION:  Fanconi Anemia Type C Gene
;    NUMBER OF SEQUENCES:  73
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Richard J. Polley, Esq.
;      ADDRESSEE:  Klarquist, Sparkman, Campbell, Leigh &
;      ADDRESSEE:  Whinston, LLP
;      STREET:  121 S.W. Salmon, Suite 1600
;      CITY:  Portland
;      STATE:  Oregon
;      COUNTRY:  U.S.A.
;      ZIP:  97204
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Disk, 3+-inch
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS DOS
;      SOFTWARE:  WordPerfect 5.1/ASCII Text File
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/441,430
;      FILING DATE:  May 15, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  U.S. 07/876,285
;      FILING DATE:  April 29, 1992
;      APPLICATION NUMBER:  U.S. 07/918,313
;      FILING DATE:  July 21, 1992
;      APPLICATION NUMBER:  U.S. 08/003,963
;      FILING DATE:  January 15, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Richard J. Polley, Esq.
;      REGISTRATION NUMBER:  28,107
;      REFERENCE/DOCKET NUMBER:  3812-42824
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (503) 226-7391
;      TELEFAX:  (503) 228-9446
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4488 base pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Double stranded
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  cDNA to mRNA
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    IMMEDIATE SOURCE:
;      LIBRARY:  Human cDNA
;    POSITION IN GENOME:  (of corresponding genomic gene)
;      CHROMOSOME/SEGMENT:  9q
;      MAP POSITION:  22.3
;      UNITS:
US-08-441-430-1


  Query Match             0.6%;  Score 17;  DB 1;  Length 4488;
  Best Local Similarity   100.0%;  Pred. No. 73;
  Matches   17;  Conservative    0;  Mismatches    0;  Indexels    0;  Gaps    0;

Qy   2105 tgaagccaccctggaag 2121
          |||||||||||||||||
Db   4189 TGAAGCCACCCTGGAAG 4205


RESULT  30
US-07-807-043B-8
;  Sequence 8, Application US/07807043B
;  Patent No. 5342774
;  GENERAL INFORMATION:
;    APPLICANT:  Boon, Thierry, Van den Eynde, Beno t
;    TITLE OF INVENTION:  Tumor Rejection Antigen Precursors, Tumor
;    TITLE OF INVENTION:  Rejection Antigens and Uses Thereof
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/807,043B
;      FILING DATE:  19911212
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/764,364
;      FILING DATE:  23-SEPTEMBER-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/728,838
;      FILING DATE:  9-JULY-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/705,702
;      FILING DATE:  23-May-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5342774man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  LUD 253.3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5674 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  singular
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  genomic DNA
;    FEATURE:
;      NAME/KEY:  MAGE-1 gene
US-07-807-043B-8
```

Query Match 0.6%; Score 17; DB 1; Length 5674;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1652 ggtcctgggcaccctgg 1668
||||||||||||||||
Db 3994 GGTCCTGGGCACCCTGG 4010

RESULT 31
US-08-190-411A-1
Sequence 1, Application US/08190411A
Patent No. 5541104
GENERAL INFORMATION:
APPLICANT: Chen, Yao-Tseng; Stockert, Elisabeth;
APPLICANT: Chen, Yachi; Garin-Chesa, Pilar; Rettig, Wolfgang J.;
APPLICANT: van der Bruggen, Pierre; Boon-Falleur, Thierry;
APPLICANT: Old, Lloyd J.
TITLE OF INVENTION: MONOCLONAL ANTIBODIES WHICH BIND TO
TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PRECURSOR MAGE-1, RECOMBINANT MAGE-1,
TITLE OF INVENTION: AND MAGE-1 DERIVED IMMUNOGENIC PEPTIDES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,411A
FILING DATE: 01-FEBRUARY-1994
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 037,230
FILING DATE: 26-MARCH-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/04354
FILING DATE: 22-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/807,043
FILING DATE: 12-DECEMBER-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/764,364
FILING DATE: 23-SEPTEMBER-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/728,838
APPLICATION NUMBER: 9-JULY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/705,702
FILING DATE: 23-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5541104man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5354
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5674 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
FEATURE:
NAME/KEY: MAGE-1 gene
US-08-190-411A-1

Query Match 0.6%; Score 17; DB 1; Length 5674;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1652 ggtcctgggcaccctgg 1668
||||||||||||||||
Db 3994 GGTCCTGGGCACCCTGG 4010

RESULT 32
US-08-299-849B-8
; Sequence 8, Application US/08299849B
; Patent No. 5612201
; GENERAL INFORMATION:
; APPLICANT: De Plaen, Etienne; Boon-Falleur, Thierry;
; APPLICANT: Leth , Bernard; Szikora, Jean-Pierre; De Smet, Charles;
; APPLICANT: Chomez, Patrick
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Useful In
; TITLE OF INVENTION: Determining Expression Of A Tumor Antigen Precursor
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/299,849B
; FILING DATE: 1-SEPTEMBER-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/037,230
; FILING DATE: 26-MARCH-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/04354
; FILING DATE: 22-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/807,043
; FILING DATE: 12-DECEMBER-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/764,364
; FILING DATE: 23-SEPTEMBER-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/728,838
; APPLICATION NUMBER: 9-JULY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/705,702
; FILING DATE: 23-May-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5612201man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5355
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5674 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; FEATURE:
; NAME/KEY: MAGE-1 gene
US-08-299-849B-8

```
  Query Match              0.6%;  Score 17;  DB 1;  Length 5674;
  Best Local Similarity   100.0%;  Pred. No. 73;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1652 ggtcctgggcaccctgg 1668
          |||||||||||||||||
Db   3994 GGTCCTGGGCACCCTGG 4010


RESULT  33
US-08-560-024-1
 Sequence 1, Application US/08560024
 Patent No. 5843448
  GENERAL INFORMATION:
    APPLICANT:  Chen, Yao-Tseng; Stockert, Elisabeth;
    APPLICANT:  Chen, Yachi; Garin-Chesa, Pilar; Rettig, Wolfgang J.;
    APPLICANT:  van der Bruggen, Pierre;  Boon-Falleur, Thierry;
    APPLICANT:  Old, Lloyd J.
    TITLE OF INVENTION:  MONOCLONAL ANTIBODIES WHICH BIND TO
    TITLE OF INVENTION:  TUMOR REJECTION ANTIGEN PRECURSOR MAGE-1, RECOMBINANT MAGE-1,
    TITLE OF INVENTION:  AND MAGE-1 DERIVED IMMUNOGENIC PEPTIDES
    NUMBER OF SEQUENCES:  4
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Felfe & Lynch
      STREET:  805 Third Avenue
      CITY:  New York City
      STATE:  New York
      ZIP:  10022
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
      COMPUTER:  IBM
      OPERATING SYSTEM:  PC-DOS
      SOFTWARE:  Wordperfect
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/560,024
      FILING DATE:
      CLASSIFICATION:  514
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  US/08/190,411
      FILING DATE:  01-FEBRUARY-1994
      APPLICATION NUMBER:  037,230
      FILING DATE:  26-MARCH-1993
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  PCT/US92/04354
      FILING DATE:  22-MAY-1992
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  07/807,043
      FILING DATE:  12-DECEMBER-1991
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  07/764,364
      FILING DATE:  23-SEPTEMBER-1991
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  07/728,838
      APPLICATION NUMBER:  9-JULY-1991
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  07/705,702
      FILING DATE:  23-MAY-1991
    ATTORNEY/AGENT INFORMATION:
      NAME:  Hanson, No. 5843448man D.
      REGISTRATION NUMBER:  30,946
      REFERENCE/DOCKET NUMBER:  LUD 5354
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (212) 688-9200
      TELEFAX:  (212) 838-3884
  INFORMATION FOR SEQ ID NO:  1:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  5674 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    MOLECULE TYPE:  genomic DNA
;    FEATURE:
;      NAME/KEY:  MAGE-1 gene
US-08-560-024-1


  Query Match              0.6%;  Score 17;  DB 2;  Length 5674;
  Best Local Similarity   100.0%;  Pred. No. 73;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1652 ggtcctgggcaccctgg 1668
          |||||||||||||||||
Db   3994 GGTCCTGGGCACCCTGG 4010


RESULT  34
US-08-142-368A-8
; Sequence 8, Application US/08142368A
; Patent No. 5925729
;  GENERAL INFORMATION:
;    APPLICANT:  Boon-Falleur, Thierry; Van der Bruggen, Thierry;
;    APPLICANT:  Van den Eynde, Beno t; Van Pel, Aline; De Plaen, Etienne;
;    APPLICANT:  Lurquin, Christophe; Chomez, Patrick; Traversari, Catia
;    TITLE OF INVENTION:  Tumor Rejection Antigen Precursors, Tumor
;    TITLE OF INVENTION:  Rejection Antigens and Uses Thereof
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/142,368A
;      FILING DATE:  02-MAY-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/04354
;      FILING DATE:  22-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/807,043
;      FILING DATE:  12-DECEMBER-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/764,364
;      FILING DATE:  23-SEPTEMBER-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/728,838
;      APPLICATION NUMBER:  9-JULY-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/705,702
;      FILING DATE:  23-May-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5925729man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  LUD 5253.4-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5674 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  genomic DNA
;    FEATURE:
;      NAME/KEY:  MAGE-1 gene
US-08-142-368A-8
```

Query Match 0.6%; Score 17; DB 2; Length 5674;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1652 ggtcctgggcaccctgg 1668
||||||||||||||||
Db 3994 GGTCCTGGGCACCCTGG 4010

RESULT 35
US-08-967-727-8
Sequence 8, Application US/08967727
Patent No. 6025474
GENERAL INFORMATION:
APPLICANT: Gaugler, B atrice; Van den Eynde, Beno t;
APPLICANT: van der Bruggen, Pierre; Boon-Falleur, Thierry
TITLE OF INVENTION: Isolated Nucleic Acid Molecules Coding For
TITLE OF INVENTION: Tumor Rejection Antigen Precursor Mage-3 And Uses Thereof
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,727
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/037,230
FILING DATE: 26-MARCH-1993
APPLICATION NUMBER: PCT/US92/04354
FILING DATE: 22-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/807,043
FILING DATE: 12-DECEMBER-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/764,365
FILING DATE: 23-SEPTEMBER-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/728,838
FILING DATE: 9-JULY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/705,702
FILING DATE: 23-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6025474man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5353
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 5674 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
FEATURE:
NAME/KEY: MAGE-1 gene
US-08-967-727-8

Query Match 0.6%; Score 17; DB 3; Length 5674;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1652 ggtcctgggcaccctgg 1668
||||||||||||||||
Db 3994 GGTCCTGGGCACCCTGG 4010

RESULT 36
US-08-321-478-6
; Sequence 6, Application US/08321478
; Patent No. 5527677
; GENERAL INFORMATION:
; APPLICANT: DEGUCHI, Takeo
; APPLICANT: KINOSHITA, Moritoshi
; APPLICANT: KATSURAGI, Kiyonori
; APPLICANT: SHIN, Sadahito
; TITLE OF INVENTION: HUMAN ARYLAMINE N-ACETYLTRANSFERASE
; TITLE OF INVENTION: GENES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/321,478
; FILING DATE: 11-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/038,667
; FILING DATE: 23-MAR-1993
; APPLICATION NUMBER: JP 64669/1992
; FILING DATE: 23-MAR-1992
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6464 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 723..1595
; FEATURE:
; NAME/KEY: exon
; LOCATION: 717..1936
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 1794..1799
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 1800..1805
US-08-321-478-6

Query Match 0.6%; Score 17; DB 1; Length 6464;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1032 aggtaccaaggaaaggc 1048
||||||||||||||||
Db 5916 AGGTACCAAGGAAAGGC 5932

```
RESULT  37
US-07-853-913-1
  Sequence 1, Application US/07853913
  Patent No. 5338839
  GENERAL INFORMATION:
    APPLICANT:  McKay, Ronald D.G.
    APPLICANT:  Lendahl, Urban
    TITLE OF INVENTION:  Nestin Expression As An Indicator of
    TITLE OF INVENTION:  Neuroepithelial Tumors
    NUMBER OF SEQUENCES:  4
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
      STREET:  Two Militia Drive
      CITY:  Lexington
      STATE:  Massachusetts
      COUNTRY:  U.S.A.
      ZIP:  02173
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/07/853,913
      FILING DATE:  19920319
      CLASSIFICATION:  435
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  US 07/660,412
      FILING DATE:  22-FEB-1991
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  US 07/603,803
      FILING DATE:  25-OCT-1990
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  US 07/201,762
      FILING DATE:  02-JUN-1988
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  US 07/180,548
    ATTORNEY/AGENT INFORMATION:
      NAME:  Granahan, Patricia
      REGISTRATION NUMBER:  32,227
      REFERENCE/DOCKET NUMBER:  MIT-4641AAAA
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  617-861-6240
      TELEFAX:  617-861-9540
  INFORMATION FOR SEQ ID NO:  1:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  11236 base pairs
      TYPE:  NUCLEIC ACID
      STRANDEDNESS:  double
      TOPOLOGY:  linear
    MOLECULE TYPE:  DNA (genomic)
US-07-853-913-1


  Query Match             0.6%;  Score 17;  DB 1;  Length 11236;
  Best Local Similarity   100.0%;  Pred. No. 74;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1877 ccttcaggaaggggctg 1893
          |||||||||||||||||
Db   3466 CCTTCAGGAAGGGGCTG 3482


RESULT  38
US-08-781-891-207
  Sequence 207, Application US/08781891
  Patent No. 6090620
  GENERAL INFORMATION:
    APPLICANT:  Fu, Ying-Hui
    APPLICANT:  Yu, Chang-En
    APPLICANT:  Oshima, Junko
    APPLICANT:  Mulligan, John T.
    APPLICANT:  Schellenberg, Gerald D.
    TITLE OF INVENTION:  GENE AND GENE PRODUCTS RELATED TO
    TITLE OF INVENTION:  WERNER'S SYNDROME
    NUMBER OF SEQUENCES:  209
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  SEED and BERRY LLP
      STREET:  6300 Columbia Center, 701 Fifth Avenue
      CITY:  Seattle
      STATE:  Washington
      COUNTRY:  USA
      ZIP:  98104-7092
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/781,891
      FILING DATE:  27-DEC-1996
      CLASSIFICATION:  800
    ATTORNEY/AGENT INFORMATION:
      NAME:  No. 6090620tenburg Ph.D., Carol
      REGISTRATION NUMBER:  39,317
      REFERENCE/DOCKET NUMBER:  240052.419
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (206) 622-4900
      TELEFAX:  (206) 682-6031
  INFORMATION FOR SEQ ID NO:  207:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  29604 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
US-08-781-891-207


  Query Match             0.6%;  Score 17;  DB 3;  Length 29604;
  Best Local Similarity   100.0%;  Pred. No. 74;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    525 aaaggaatagaactggc 541
          |||||||||||||||||
Db  14826 AAAGGAATAGAACTGGC 14842


RESULT  39
US-09-335-409-1/c
  Sequence 1, Application US/09335409
  Patent No. 6121029
  GENERAL INFORMATION:
   APPLICANT: Schupp, Thomas
   APPLICANT: Ligon, James
   APPLICANT: Molnar, Istvan
   APPLICANT: Zirkle, Ross
   APPLICANT: Cyr, Devon
   APPLICANT: Goerlach, Joern
   TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
   FILE REFERENCE: 4-30582A
   CURRENT APPLICATION NUMBER: US/09/335,409
   CURRENT FILING DATE: 1999-06-17
   NUMBER OF SEQ ID NOS: 30
   SOFTWARE: PatentIn Ver. 2.0
  SEQ ID NO 1
    LENGTH: 68750
    TYPE: DNA
    ORGANISM: Sorangium cellulosum
US-09-335-409-1


  Query Match             0.6%;  Score 17;  DB 3;  Length 68750;
  Best Local Similarity   100.0%;  Pred. No. 75;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy   1457 cggcccagccccagcag 1473
          ||||||||||||||||
Db  20332 CGGCCCAGCCCCAGCAG 20316


RESULT  40
US-07-753-110B-12
; Sequence 12, Application US/07753110B
; Patent No. 5436141
;  GENERAL INFORMATION:
;    APPLICANT:  Miyata, Shohei
;    APPLICANT:  Ohshima, Atsushi
;    APPLICANT:  Inouye, Sumiko
;    APPLICANT:  Inouye, Masayori
;    TITLE OF INVENTION:  METHOD FOR SYNTHESIZING STABLE
;    TITLE OF INVENTION:  SINGLE-STRANDED cDNA IN EUKARYOTES BY MEANS OF A BACTERIAL
;    TITLE OF INVENTION:  RETRON, PRODUCTS AND USES THEREFOR
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Weiser & Associates
;      STREET:  230 South Fifteenth Street, Suite 500
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/753,110B
;      FILING DATE:  30-AUG-1991
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weiser, Gerard J.
;      REGISTRATION NUMBER:  19,763
;      REFERENCE/DOCKET NUMBER:  377.5584P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-875-8383
;      TELEFAX:  215-875-8394
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  76 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  19
;      OTHER INFORMATION:  /note= "The 2' position of this
;      OTHER INFORMATION:  nucleotide is linked to the 5' position of
;      OTHER INFORMATION:  nucleotide number 1 of SEQ ID NO: 11 of this
;      OTHER INFORMATION:  application."
;    FEATURE:
;      NAME/KEY:  misc_binding
;      LOCATION:  69..76
;      OTHER INFORMATION:  /note= "This region can hydrogen
;      OTHER INFORMATION:  bond to nucleotides 156-163 of SEQ ID NO: 11 of
;      OTHER INFORMATION:  this application."
US-07-753-110B-12


  Query Match             0.5%;  Score 16;  DB 1;  Length 76;
  Best Local Similarity   87.5%;  Pred. No. 2.2e+02;
  Matches   14;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2165 gtcccaagccatcagc 2180
          |:||||||||||:||||
Db      5 GUCCCAAGCCAUCAGC 20


RESULT  41
US-08-503-730-16
; Sequence 16, Application US/08503730
; Patent No. 5780269
;  GENERAL INFORMATION:
;    APPLICANT:  Inouye, Sumiko
;    APPLICANT:  Inouye, Masayori
;    TITLE OF INVENTION:  NEW HYBRID MOLECULES
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Weiser & Associates
;      STREET:  230 South Fifteenth Street Suite 500
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/503,730
;      FILING DATE:  18-JUL-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/817,430
;      FILING DATE:  06-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weiser, Gerard J.
;      REGISTRATION NUMBER:  19,763
;      REFERENCE/DOCKET NUMBER:  377(913).6277P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-875-8383
;      TELEFAX:  215-875-8394
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  76 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  both
US-08-503-730-16


  Query Match             0.5%;  Score 16;  DB 1;  Length 76;
  Best Local Similarity   87.5%;  Pred. No. 2.2e+02;
  Matches   14;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2165 gtcccaagccatcagc 2180
          |:||||||||||:||||
Db      5 GUCCCAAGCCAUCAGC 20


RESULT  42
US-08-507-634-13
; Sequence 13, Application US/08507634
; Patent No. 5849563
;  GENERAL INFORMATION:
;    APPLICANT:  Miyata, Shohei
;    APPLICANT:  Ohshima, Atsushi
;    APPLICANT:  Inouye, Sumiko
;    APPLICANT:  Inouye, Masayori
;    TITLE OF INVENTION:  METHOD FOR SYNTHESIZING STABLE
;    TITLE OF INVENTION:  SINGLE-STRANDED cDNA IN EUKARYOTES BY MEANS OF A BACTERIAL
;    TITLE OF INVENTION:  RETRON, PRODUCTS AND USES THEREFOR
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Weiser & Associates
;      STREET:  230 South Fifteenth Street, Suite 500
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19102
```

```
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/507,634
      FILING DATE:  25-JUL-1995
      CLASSIFICATION:  435
    ATTORNEY/AGENT INFORMATION:
      NAME:  Weiser, Gerard J.
      REGISTRATION NUMBER:  19,763
      REFERENCE/DOCKET NUMBER:  377.6282P
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  215-875-8383
      TELEFAX:  215-875-8394
  INFORMATION FOR SEQ ID NO:  13:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  76 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    FEATURE:
      NAME/KEY:  misc_feature
      LOCATION:  19
      OTHER INFORMATION:  /note= "The 2' position of this
      OTHER INFORMATION:  nucleotide is linked to the 5' position of
      OTHER INFORMATION:  nucleotide number 1 of SEQ ID NO: 12 of this
      OTHER INFORMATION:  application."
    FEATURE:
      NAME/KEY:  misc_binding
      LOCATION:  69..76
      OTHER INFORMATION:  /note= "This region can hydrogen
      OTHER INFORMATION:  bond to nucleotides 156-163 of SEQ ID NO: 12 of
      OTHER INFORMATION:  this application."
US-08-507-634-13


  Query Match             0.5%;  Score 16;  DB 2;  Length 76;
  Best Local Similarity   87.5%;  Pred. No. 2.2e+02;
  Matches   14;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2165 gtcccaagccatcagc 2180
          |:||||||||:||||
Db      5 GUCCCAAGCCAUCAGC 20


RESULT  43
US-08-717-294-93/c
; Sequence 93, Application US/08717294
; Patent No. 6114148
  GENERAL INFORMATION:
    APPLICANT:  SEED, BRIAN
    APPLICANT:  HAAS, JURGEN
    TITLE OF INVENTION:  HIGH LEVEL EXPRESSION OF
    TITLE OF INVENTION:  PROTEINS
    NUMBER OF SEQUENCES:  110
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Clark & Elbing LLP
      STREET:  176 Federal Street
      CITY:  Boston
      STATE:  MA
      COUNTRY:  USA
      ZIP:  02110
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Diskette
      COMPUTER:  IBM Compatible
      OPERATING SYSTEM:  DOS
      SOFTWARE:  FastSEQ for Windows Version 2.0
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/717,294
      FILING DATE:  20-SEP-1996
      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Elbing, Karen L
;      REGISTRATION NUMBER:  35,238
;      REFERENCE/DOCKET NUMBER:  00786/345001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-428-0200
;      TELEFAX:  617-428-7045
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  93:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  105 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Other
US-08-717-294-93


  Query Match             0.5%;  Score 16;  DB 3;  Length 105;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1101 gacagcaggtaccagc 1116
          ||||||||||||||||
Db     72 GACAGCAGGTACCAGC 57


RESULT  44
US-08-687-080-70
; Sequence 70, Application US/08687080
; Patent No. 5965427
  GENERAL INFORMATION:
    APPLICANT:  Gregory Dolganov
    TITLE OF INVENTION:  Human RAD50 Gene and Methods of Use Thereof
    NUMBER OF SEQUENCES:  175
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Dehlinger & Associates
      STREET:  350 Cambridge Avenue, Suite 250
      CITY:  Palo Alto
      STATE:  CA
      COUNTRY:  USA
      ZIP:  94306
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/687,080
      FILING DATE:  17-JUL-1996
      CLASSIFICATION:  435
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  US 08/592,126
      FILING DATE:  26-JAN-1996
    ATTORNEY/AGENT INFORMATION:
      NAME:  Sholtz, Charles K.
      REGISTRATION NUMBER:  38,615
      REFERENCE/DOCKET NUMBER:  4600-0111.30
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (415) 324-0880
      TELEFAX:  (415) 324-0960
  INFORMATION FOR SEQ ID NO:  70:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  233 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  double
      TOPOLOGY:  linear
    MOLECULE TYPE:  DNA (genomic)
    HYPOTHETICAL:  NO
    ANTI-SENSE:  NO
```

```
    ORIGINAL SOURCE:
      INDIVIDUAL ISOLATE:  INTRON 6 OF RAD50 GENOMIC SEQUENCE
S-08-687-080-70


  Query Match             0.5%;  Score 16;  DB 2;  Length 233;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

y    966 tttgtggtggtagaat 981
         ||||||||||||||||
b      2 TTTGTGGTGGTAGAAT 17


ESULT  45
S-08-630-822A-97/c
 Sequence 97, Application US/08630822A
 Patent No. 5840695
 GENERAL INFORMATION:
   APPLICANT:  FRANK, GLENN R.
   APPLICANT:  HUNTER, SHIRLEY WU
   APPLICANT:  WALLENFELS, LYNDA
   TITLE OF INVENTION:  NOVEL ECTOPARASITE SALIVA PROTEINS
   TITLE OF INVENTION:  AND APPARATUS TO COLLECT SUCH PROTEINS
   NUMBER OF SEQUENCES:  107
   CORRESPONDENCE ADDRESS:
     ADDRESSEE:  Sheridan Ross P.C.
     STREET:  1700 Lincoln Street, Suite 3500
     CITY:  Denver
     STATE:  Colorado
     COUNTRY:  U.S.A.
     ZIP:  80203
   COMPUTER READABLE FORM:
     MEDIUM TYPE:  Floppy disk
     COMPUTER:  IBM PC compatible
     OPERATING SYSTEM:  PC-DOS/MS-DOS
     SOFTWARE:  PatentIn Release #1.0, Version #1.25
   CURRENT APPLICATION DATA:
     APPLICATION NUMBER:  US/08/630,822A
     FILING DATE:  11-APR-1996
     CLASSIFICATION:  435
   ATTORNEY/AGENT INFORMATION:
     NAME:  CONNELL, GARY J.
     REGISTRATION NUMBER:  32,020
     REFERENCE/DOCKET NUMBER:  2618-17-C3
   TELECOMMUNICATION INFORMATION:
     TELEPHONE:  (303) 863-9700
     TELEFAX:  (303) 863-0223
  INFORMATION FOR SEQ ID NO:  97:
   SEQUENCE CHARACTERISTICS:
     LENGTH:  252 base pairs
     TYPE:  nucleic acid
     STRANDEDNESS:  single
     TOPOLOGY:  linear
   MOLECULE TYPE:  cDNA
S-08-630-822A-97


  Query Match             0.5%;  Score 16;  DB 2;  Length 252;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1478 aagaagtcagtaccca 1493
          ||||||||||||||||
Db    195 AAGAAGTCAGTACCCA 180


RESULT  46
S-09-005-069-97/c
 Sequence 97, Application US/09005069
 Patent No. 5932470
 GENERAL INFORMATION:
   APPLICANT:  FRANK, GLENN R.
;    APPLICANT:  HUNTER, SHIRLEY WU
;    APPLICANT:  WALLENFELS, LYNDA
;    TITLE OF INVENTION:  NOVEL ECTOPARASITE SALIVA PROTEINS
;    TITLE OF INVENTION:  AND APPARATUS TO COLLECT SUCH PROTEINS
;    NUMBER OF SEQUENCES:  107
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sheridan Ross P.C.
;      STREET:  1700 Lincoln Street, Suite 3500
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  U.S.A.
;      ZIP:  80203
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/005,069
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/630,822
;      FILING DATE:  11-APR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CONNELL, GARY J.
;      REGISTRATION NUMBER:  32,020
;      REFERENCE/DOCKET NUMBER:  2618-17-C3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 863-9700
;      TELEFAX:  (303) 863-0223
;  INFORMATION FOR SEQ ID NO:  97:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  252 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-005-069-97


  Query Match             0.5%;  Score 16;  DB 2;  Length 252;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1478 aagaagtcagtaccca 1493
          ||||||||||||||||
Db    195 AAGAAGTCAGTACCCA 180


RESULT  47
US-08-906-769-104/c
; Sequence 104, Application US/08906769
; Patent No. 6077687
;  GENERAL INFORMATION:
;    APPLICANT:  Grieve, Robert B.
;    APPLICANT:  Rushlow, Keith E.
;    APPLICANT:  Wu Hunter, Shirley
;    APPLICANT:  Frank, Glenn R.
;    APPLICANT:  Stiegler, Gary
;    APPLICANT:  Gaines, Patrick J.
;    APPLICANT:  Silver, Gary
;    TITLE OF INVENTION:  FLEA PROTEASE PROTEINS, NUCLEIC ACID
;    TITLE OF INVENTION:  MOLECULES AND USES THEREOF
;    NUMBER OF SEQUENCES:  190
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sheridan Ross & McIntosh
;      STREET:  1700 Lincoln Street, Suite 3500
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80203
;  COMPUTER READABLE FORM:
```

```
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/906,769
      FILING DATE:
      CLASSIFICATION:
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  US 08/639,075
      FILING DATE:  24-APR-1996
    ATTORNEY/AGENT INFORMATION:
      NAME:  Connell, Gary J.
      REGISTRATION NUMBER:  32,020
      REFERENCE/DOCKET NUMBER:  2618-25-C2
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (303) 863-9700
      TELEFAX:  (303) 863-0223
  INFORMATION FOR SEQ ID NO:  104:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  252 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    MOLECULE TYPE:  cDNA
    FEATURE:
      NAME/KEY:  CDS
      LOCATION:  1..251
      OTHER INFORMATION:  /note= "At pos. bp 4, change A to
      OTHER INFORMATION:  R.  At pos. aa 2, substitute Xaa."
S-08-906-769-104


  Query Match              0.5%;  Score 16;  DB 3;  Length 252;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1478 aagaagtcagtaccca 1493
          ||||||||||||||||
Db    199 AAGAAGTCAGTACCCA 184


RESULT  48
S-08-906-616-104/c
 Sequence 104, Application US/08906616
 Patent No. 6121035
  GENERAL INFORMATION:
    APPLICANT:  Grieve, Robert B.
    APPLICANT:  Rushlow, Keith E.
    APPLICANT:  Wu Hunter, Shirley
    APPLICANT:  Frank, Glenn R.
    APPLICANT:  Stiegler, Gary
    APPLICANT:  Gaines, Patrick J.
    APPLICANT:  Silver, Gary
    TITLE OF INVENTION:  FLEA AMINOPEPTIDASE PROTEINS AND USES THEREOF
    NUMBER OF SEQUENCES:  190
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Sheridan Ross P.C.
      STREET:  1700 Lincoln Street, Suite 3500
      CITY:  Denver
      STATE:  Colorado
      COUNTRY:  USA
      ZIP:  80203
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/906,616
      FILING DATE:  05-AUG-1997
      CLASSIFICATION:  536
    ATTORNEY/AGENT INFORMATION:
      NAME:  Connell, Gary J.
      REGISTRATION NUMBER:  32,020
      REFERENCE/DOCKET NUMBER:  2618-25-C2-3
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (303) 863-9700
      TELEFAX:  (303) 863-0223
  INFORMATION FOR SEQ ID NO:  104:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  252 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    MOLECULE TYPE:  cDNA
    FEATURE:
      NAME/KEY:  CDS
      LOCATION:  1..251
      OTHER INFORMATION:  /note= "At pos. bp 4, change A to
      OTHER INFORMATION:  R.  At pos. aa 2, substitute Xaa."
US-08-906-616-104


  Query Match              0.5%;  Score 16;  DB 3;  Length 252;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1478 aagaagtcagtaccca 1493
          ||||||||||||||||
Db    199 AAGAAGTCAGTACCCA 184


RESULT  49
US-08-817-795-104/c
 Sequence 104, Application US/08817795
 Patent No. 6139840
  GENERAL INFORMATION:
    APPLICANT:  Grieve, Robert B.
    APPLICANT:  Rushlow, Keith E.
    APPLICANT:  Hunter, Shirley Wu
    APPLICANT:  Frank, Glenn R.
    APPLICANT:  Heath, Andrew W.
    APPLICANT:  Yamaka, Miles Yamanaka
    APPLICANT:  Arfsten, Ann
    APPLICANT:  Dale, Beverly
    APPLICANT:  Stiegler, Gary
    TITLE OF INVENTION:  USE OF PROTEASE INHIBITORS AND
    TITLE OF INVENTION:  PROTEASE VACCINES TO PROTECT  ANIMALS FROM FLEA
    TITLE OF INVENTION:  INFESTATION, AND FLEA PROTEASE PROTEINS, NUCLEIC ACID
    TITLE OF INVENTION:  MOLECULES, AND USES THEREOF
    NUMBER OF SEQUENCES:  119
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Sheridan Ross & McIntosh
      STREET:  1700 Lincoln Street, Suite 3500
      CITY:  Denver
      STATE:  Colorado
      COUNTRY:  USA
      ZIP:  80203
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/817,795
      FILING DATE:
      CLASSIFICATION:
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  PCT/US95/14442
      FILING DATE:
    ATTORNEY/AGENT INFORMATION:
      NAME:  Gary J. Connell
      REGISTRATION NUMBER:  32,020
      REFERENCE/DOCKET NUMBER:
    TELECOMMUNICATION INFORMATION:
```

```
;       TELEPHONE:  (303) 863-9700
;       TELEFAX:  (303 863-0223
;     INFORMATION FOR SEQ ID NO:  104:
;       SEQUENCE CHARACTERISTICS:
;         LENGTH:  252 base pairs
;         TYPE:  nucleic acid
;         STRANDEDNESS:  single
;         TOPOLOGY:  linear
;       MOLECULE TYPE:  cDNA
;       FEATURE:
;         NAME/KEY:  CDS
;         LOCATION:  1..251
;       FEATURE:
;         NAME/KEY:  Xaa = any amino acid
;         LOCATION:  2
US-08-817-795-104


  Query Match             0.5%;  Score 16;  DB 3;  Length 252;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1478 aagaagtcagtaccca 1493
          ||||||||||||||||
Db    199 AAGAAGTCAGTACCCA 184


RESULT  50
US-08-639-075A-104/c
 Sequence 104, Application US/08639075A
 Patent No. 6150125
  GENERAL INFORMATION:
    APPLICANT:  Grieve, Robert B.
    APPLICANT:  Rushlow, Keith E.
    APPLICANT:  Wu Hunter, Shirley
    APPLICANT:  Frank, Glenn R.
    APPLICANT:  Stiegler, Gary
    APPLICANT:  Gaines, Patrick J.
    APPLICANT:  Silver, Gary
    TITLE OF INVENTION:  FLEA PROTEASE PROTEINS, NUCLEIC ACID
    TITLE OF INVENTION:  MOLECULES AND USES THEREOF
    NUMBER OF SEQUENCES:  190
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Sheridan Ross & McIntosh
      STREET:  1700 Lincoln Street, Suite 3500
      CITY:  Denver
      STATE:  Colorado
      COUNTRY:  USA
      ZIP:  80203
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/639,075A
      FILING DATE:  24-APR-1996
      CLASSIFICATION:  424
    ATTORNEY/AGENT INFORMATION:
      NAME:  Connell, Gary J.
      REGISTRATION NUMBER:  32,020
      REFERENCE/DOCKET NUMBER:  2618-25-C2
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (303) 863-9700
      TELEFAX:  (303) 863-0223
  INFORMATION FOR SEQ ID NO:  104:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  252 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    MOLECULE TYPE:  cDNA
    FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  1..251
;       OTHER INFORMATION:  /note= "At pos. bp 4, change A to
;       OTHER INFORMATION:  R.  At pos. aa 2, substitute Xaa."
US-08-639-075A-104


  Query Match             0.5%;  Score 16;  DB 3;  Length 252;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1478 aagaagtcagtaccca 1493
          ||||||||||||||||
Db    199 AAGAAGTCAGTACCCA 184


RESULT  51
PCT-US95-14442A-104/c
; Sequence 104, Application PC/TUS9514442A
;  GENERAL INFORMATION:
;    APPLICANT:  Grieve, Robert B.
;    APPLICANT:  Rushlow, Keith E.
;    APPLICANT:  Hunter, Shirley Wu
;    APPLICANT:  Frank, Glenn R.
;    APPLICANT:  Heath, Andrew W.
;    APPLICANT:  Yamaka, Miles Yamanaka
;    APPLICANT:  Arfsten, Ann
;    APPLICANT:  Dale, Beverly
;    APPLICANT:  Stiegler, Gary
;    TITLE OF INVENTION:  USE OF PROTEASE INHIBITORS AND
;    TITLE OF INVENTION:  PROTEASE VACCINES TO PROTECT  ANIMALS FROM FLEA
;    TITLE OF INVENTION:  INFESTATION, AND FLEA PROTEASE PROTEINS, NUCLEIC ACID
;    TITLE OF INVENTION:  MOLECULES, AND USES THEREOF
;    NUMBER OF SEQUENCES:  119
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sheridan Ross & McIntosh
;      STREET:  1700 Lincoln Street, Suite 3500
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80203
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/14442A
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gary J. Connell
;      REGISTRATION NUMBER:  32,020
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 863-9700
;      TELEFAX:  (303 863-0223
;  INFORMATION FOR SEQ ID NO:  104:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  252 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..251
;    FEATURE:
;      NAME/KEY:  Xaa = any amino acid
;      LOCATION:  2
PCT-US95-14442A-104
```

```
  Query Match             0.5%;  Score 16;  DB 4;  Length 252;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1478 aagaagtcagtaccca 1493
          ||||||||||||||||
Db    199 AAGAAGTCAGTACCCA 184


RESULT  52
US-08-611-757-98/c
; Sequence 98, Application US/08611757
; Patent No. 5859230
  GENERAL INFORMATION:
    APPLICANT:  Kim, Jungsuh P.
    APPLICANT:  Reyes, Gregory R.
    APPLICANT:  Wages, John
    APPLICANT:  Zhang-Keck, Zhen-Yang
    APPLICANT:  Young, LaVonne
    TITLE OF INVENTION:  No. 5859230-A/No. 5859230-B/No. 5859230-C/No. 5859230-D/No. 58
    TITLE OF INVENTION:  Agents and Molecular Cloning Thereof
    NUMBER OF SEQUENCES:  106
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Dehlinger & Associates
      STREET:  350 Cambridge Avenue, Suite 250
      CITY:  Palo Alto
      STATE:  CA
      COUNTRY:  USA
      ZIP:  94306
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/611,757
      FILING DATE:
      CLASSIFICATION:  435
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  08/246,985
      FILING DATE:  20-MAY-1994
      APPLICATION NUMBER:  US 025,396
      FILING DATE:  24-FEB-1993
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  US 07/922,493
      FILING DATE:  30-JUL-1992
    ATTORNEY/AGENT INFORMATION:
      NAME:  Fabian, Gary R.
      REGISTRATION NUMBER:  33,875
      REFERENCE/DOCKET NUMBER:  4600-0201
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (415) 324-0880
      TELEFAX:  (415) 324-0960
  INFORMATION FOR SEQ ID NO:  98:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  294 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  double
      TOPOLOGY:  linear
    MOLECULE TYPE:  DNA
    HYPOTHETICAL:  NO
    ANTI-SENSE:  NO
    ORIGINAL SOURCE:
      INDIVIDUAL ISOLATE:  SCH Clone SU7-8
US-08-611-757-98


  Query Match             0.5%;  Score 16;  DB 2;  Length 294;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    133 gccgcgagcggccgcg 148
          ||||||||||||||||
Db     21 GCCGCGAGCGGCCGCG 6


RESULT  53
PCT-US95-05980-98/c
; Sequence 98, Application PC/TUS9505980
;  GENERAL INFORMATION:
;    APPLICANT:
;    APPLICANT:
;    TITLE OF INVENTION:  Non-A/Non-B/Non-C/Non-D/Non-E Hepatitis
;    TITLE OF INVENTION:  Agents and Molecular Cloning Thereof
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  350 Cambridge Avenue, Suite 250
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/05980
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/246,986
;      FILING DATE:  20-MAY-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0201.49
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  98:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  294 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  SCH Clone SU7-8
PCT-US95-05980-98


  Query Match             0.5%;  Score 16;  DB 4;  Length 294;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    133 gccgcgagcggccgcg 148
          ||||||||||||||||
Db     21 GCCGCGAGCGGCCGCG 6


RESULT  54
US-08-937-931-9
; Sequence 9, Application US/08937931
; Patent No. 5935792
;  GENERAL INFORMATION:
;    APPLICANT:  Rubin, Gerald M.
;    APPLICANT:  Pan, Duojia
;    APPLICANT:  Rooke, Jenny
;    APPLICANT:  Yavari, Reza
;    APPLICANT:  Xu, Tian
;    TITLE OF INVENTION:  KUZ: A No. 5935792el Family of Metalloproteases
;    NUMBER OF SEQUENCES:  10
```

```
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  SCIENCE & TECHNOLOGY LAW GROUP
      STREET:  268 BUSH STREET, SUITE 3200
      CITY:  SAN FRANCISCO
      STATE:  CALIFORNIA
      COUNTRY:  USA
      ZIP:  94104
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/937,931
      FILING DATE:
      CLASSIFICATION:  800
    ATTORNEY/AGENT INFORMATION:
      NAME:  OSMAN, RICHARD A
      REGISTRATION NUMBER:  36,627
      REFERENCE/DOCKET NUMBER:  B97-081
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (415) 343-4341
      TELEFAX:  (415) 343-4342
  INFORMATION FOR SEQ ID NO:  9:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  486 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  double
      TOPOLOGY:  linear
    MOLECULE TYPE:  cDNA
S-08-937-931-9


  Query Match             0.5%;  Score 16;  DB 2;  Length 486;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    426 cttactttaaaggaaa 441
          ||||||||||||||||
Db    428 CTTACTTTAAAGGAAA 443


RESULT  55
S-08-334-254-7
 Sequence 7, Application US/08334254
 Patent No. 5723290
  GENERAL INFORMATION:
    APPLICANT:  James Eberwine, Marc Dichter, Kevin Miyashiro
    TITLE OF INVENTION:  USE OF NEURITE LOCALIZED mRNAs FOR
    TITLE OF INVENTION:  MEDICAL DIAGNOSIS AND THERAPEUTICS
    NUMBER OF SEQUENCES:  28
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Jane Massey Licata, Esq.
      STREET:  210 Lake Drive East, Suite 201
      CITY:  Cherry Hill
      STATE:  NJ
      COUNTRY:  USA
      ZIP:  08002
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
      COMPUTER:  IBM 486
      OPERATING SYSTEM:  WINDOWS FOR WORKGROUPS
      SOFTWARE:  WORDPERFECT 5.1
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/334,254
      FILING DATE:  Herewith
      CLASSIFICATION:  436
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:
      FILING DATE:
    ATTORNEY/AGENT INFORMATION:
      NAME:  Jane Massey Licata
      REGISTRATION NUMBER:  32,257
      REFERENCE/DOCKET NUMBER:  PENN-0028
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (609) 779-2400
      TELEFAX:  (609) 779-8488
  INFORMATION FOR SEQ ID NO:  7:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  489
      TYPE:  NUCLEIC ACID
      STRANDEDNESS:  SINGLE
      TOPOLOGY:  LINEAR
    ANTI-SENSE:  NO
US-08-334-254-7


  Query Match             0.5%;  Score 16;  DB 1;  Length 489;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1881 caggaaggggctgaga 1896
          ||||||||||||||||
Db    184 CAGGAAGGGGCTGAGA 199


RESULT  56
US-08-848-131-7
; Sequence 7, Application US/08848131
; Patent No. 5958688
;  GENERAL INFORMATION:
;    APPLICANT:  James Eberwine, Marc Dichter, Kevin Miyashiro
;    TITLE OF INVENTION:  Characterization of mRNA Patterns
;    TITLE OF INVENTION:  in Neurites and Single Cells for Medical Diagnosis and
;    TITLE OF INVENTION:  Therapeutics
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jane Massey Licata, Esq.
;      STREET:  210 Lake Drive East, Suite 201
;      CITY:  Cherry Hill
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08002
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM 486
;      OPERATING SYSTEM:  WINDOWS FOR WORKGROUPS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/848,131
;      FILING DATE:  Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/334,254
;      FILING DATE:  No. 5958688ember 3, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/14792
;      FILING DATE:  No. 5958688ember 3, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  PENN-00219
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (609) 779-2400
;      TELEFAX:  (609) 779-8488
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  489
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
;    ANTI-SENSE:  NO
US-08-848-131-7


  Query Match             0.5%;  Score 16;  DB 2;  Length 489;
```

```
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1881 caggaaggggctgaga 1896
          ||||||||||||||||
Db    184 CAGGAAGGGGCTGAGA 199


RESULT  57
PCT-US95-14792-7
; Sequence 7, Application PC/TUS9514792
;  GENERAL INFORMATION:
;    APPLICANT:  James Eberwine, Marc Dichter, Kevin Miyashiro
;    TITLE OF INVENTION:  USE OF NEURITE LOCALIZED mRNAs FOR
;    TITLE OF INVENTION:  MEDICAL DIAGNOSIS AND THERAPEUTICS
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jane Massey Licata, Esq.
;      STREET:  210 Lake Drive East, Suite 201
;      CITY:  Cherry Hill
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08002
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM 486
;      OPERATING SYSTEM:  WINDOWS FOR WORKGROUPS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/14792
;      FILING DATE:  Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  PENN-0028
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (609) 779-2400
;      TELEFAX:  (609) 779-8488
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  489
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
;    ANTI-SENSE:  NO
PCT-US95-14792-7


  Query Match             0.5%;  Score 16;  DB 4;  Length 489;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1881 caggaaggggctgaga 1896
          ||||||||||||||||
Db    184 CAGGAAGGGGCTGAGA 199


RESULT  58
US-08-507-016-8
; Sequence 8, Application US/08507016
; Patent No. 5756460
;  GENERAL INFORMATION:
;    APPLICANT:  EVANS, HELEN F.
;    APPLICANT:  SHINE, JOHN
;    TITLE OF INVENTION:  HUMAN GALANIN, cDNA CLONES ENCODING
;    TITLE OF INVENTION:  HUMAN GALANIN AND A METHOD OF PRODUCING HUMAN GALANIN
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ROTHWELL, FIGG, ERNST & KURZ
;      STREET:  555 THIRTEENTH STREET, N.W.
;      CITY:  WASHINGTON
;      STATE:  D. C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/507,016
;      FILING DATE:  25-JULY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/108,733
;      FILING DATE:  03-SEP-1993
;      APPLICATION NUMBER:  PCT/AU92/00097
;      FILING DATE:  06-MAR-1992
;      APPLICATION NUMBER:  AU PK4953
;      FILING DATE:  06-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ERNST, BARBARA G.
;      REGISTRATION NUMBER:  30,377
;      REFERENCE/DOCKET NUMBER:  1871-117A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)783-6040
;      TELEFAX:  (202)783-6031
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  575 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    HYPOTHETICAL:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  14..385
US-08-507-016-8


  Query Match             0.5%;  Score 16;  DB 1;  Length 575;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2400 aagcgggagctgcggc 2415
          ||||||||||||||||
Db    200 AAGCGGGAGCTGCGGC 215


RESULT  59
PCT-US91-06418-4/c
; Sequence 4, Application PC/TUS9106418
;  GENERAL INFORMATION:
;    APPLICANT:  Oklahoma Medical Research, Foundation, et al
;    TITLE OF INVENTION:  Antigens Associated with Polymyositis
;    TITLE OF INVENTION:  and with Dermatomyositis
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kilpatrick & Cody
;      STREET:  100 Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  US
;      ZIP:  30303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/06418
```

FILING DATE: 19910905
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/579023
FILING DATE: 09-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRF120
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-572-6508
TELEFAX: 404-572-6555
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 578 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Homo sapien
STRAIN: JH2
TISSUE TYPE: Sera
CT-US91-06418-4

Query Match 0.5%; Score 16; DB 4; Length 578;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 955 ctggtgctgcttttgt 970
||||||||||||||||
Db 22 CTGGTGCTGCTTTTGT 7

RESULT 60
US-08-338-579A-94/c
Sequence 94, Application US/08338579A
Patent No. 6068975
GENERAL INFORMATION:
APPLICANT: Gilliam, T. Conrad
APPLICANT: Tanzi, Rudolph E.
TITLE OF INVENTION: ISOLATION AND USES OF A WILSON'S
TITLE OF INVENTION: DISEASE GENE
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/338,579A
FILING DATE: June 17, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/44011-A-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX:
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 609 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 163..609
US-08-338-579A-94

Query Match 0.5%; Score 16; DB 3; Length 609;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2740 gggccaggaggctgcc 2755
||||||||||||||||
Db 564 GGGCCAGGAGGCTGCC 549

RESULT 61
US-08-911-319A-2
Sequence 2, Application US/08911319A
Patent No. 5968798
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: HUMAN GLUTAREDOXIN BETA
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,319A
FILING DATE: August 14, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Muenzen, Colette C.
REGISTRATION NUMBER: 39,784
REFERENCE/DOCKET NUMBER: PF-0363 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 654 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: THP1NOT03
CLONE: 2447829
US-08-911-319A-2

Query Match 0.5%; Score 16; DB 2; Length 654;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2402 gcgggagctgcggcag 2417

```
          ||||||||||||||||
Db    101 GCGGGAGCTGCGGCAG 116


RESULT  62
US-09-352-619-2
; Sequence 2, Application US/09352619
; Patent No. 6084070
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  HUMAN GLUTAREDOXIN BETA
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/352,619
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/911,319
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Muenzen, Colette C.
;      REGISTRATION NUMBER:  39,784
;      REFERENCE/DOCKET NUMBER:  PF-0363 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  654 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  THP1NOT03
;      CLONE:  2447829
US-09-352-619-2


  Query Match           0.5%;  Score 16;  DB 3;  Length 654;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2402 gcgggagctgcggcag 2417
          ||||||||||||||||
Db    101 GCGGGAGCTGCGGCAG 116


RESULT  63
US-07-789-738-3
; Sequence 3, Application US/07789738
; Patent No. 5824857
;  GENERAL INFORMATION:
;    APPLICANT:  Beachy, Roger N.
;    APPLICANT:  Bhattacharyya, Maitrayee
;    TITLE OF INVENTION:  Plant Promoter
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Parkway No. 5824857th
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/789,738
;      FILING DATE:  19920330
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(10540)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-6099
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  714 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-07-789-738-3


  Query Match           0.5%;  Score 16;  DB 1;  Length 714;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1842 gtcctgcaccacatca 1857
          ||||||||||||||||
Db    484 GTCCTGCACCACATCA 499


RESULT  64
US-07-789-738-5
; Sequence 5, Application US/07789738
; Patent No. 5824857
;  GENERAL INFORMATION:
;    APPLICANT:  Beachy, Roger N.
;    APPLICANT:  Bhattacharyya, Maitrayee
;    TITLE OF INVENTION:  Plant Promoter
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Parkway No. 5824857th
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/789,738
;      FILING DATE:  19920330
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(10540)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-6099
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  773 base pairs
```

TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
S-07-789-738-5

Query Match 0.5%; Score 16; DB 1; Length 773;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1842 gtcctgcaccacatca 1857
||||||||||||||||
b 484 GTCCTGCACCACATCA 499

ESULT 65
S-08-651-136C-15
Sequence 15, Application US/08651136C
Patent No. 6001639
GENERAL INFORMATION:
APPLICANT: Schulein, Martin
APPLICANT: Andersen, Lene N.
APPLICANT: Lassen, Soren F.
APPLICANT: Kauppinen, Markus S.
APPLICANT: Lange, Lene
APPLICANT: Nielsen, Ruby I.
APPLICANT: Ihara, Michiko
APPLICANT: Takagi, Shinobu
TITLE OF INVENTION: No. 6001639el Endoglucanases
NUMBER OF SEQUENCES: 109
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6001639o No. 6001639disk of No. 6001639th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/651,136C
FILING DATE: 21-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4366.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 808 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 37..714
S-08-651-136C-15

Query Match 0.5%; Score 16; DB 3; Length 808;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 1409 tcccaacttccagcag 1424
||||||||||||||||

Db 577 TCCCAACTTCCAGCAG 592

RESULT 66
US-08-906-769-128/c
; Sequence 128, Application US/08906769
; Patent No. 6077687
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Silver, Gary
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,769
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/639,075
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 128:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 815 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..762
; OTHER INFORMATION: /note= "At pos. bp 453, change A to
; OTHER INFORMATION: M; at 454, change G to V; at 456, G to V; at 457, A to M;
; OTHER INFORMATION: 460, A to R; at 470, G to S; at 493, A to R. At pos. aa 1
; OTHER INFORMATION: 136, 152, 153, 154, 157 and 165, substitute Xaa."
US-08-906-769-128

Query Match 0.5%; Score 16; DB 3; Length 815;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1478 aagaagtcagtaccca 1493
||||||||||||||||
Db 181 AAGAAGTCAGTACCCA 166

RESULT 67
US-08-906-616-128/c
; Sequence 128, Application US/08906616

```
Patent No. 6121035
 GENERAL INFORMATION:
   APPLICANT:  Grieve, Robert B.
   APPLICANT:  Rushlow, Keith E.
   APPLICANT:  Wu Hunter, Shirley
   APPLICANT:  Frank, Glenn R.
   APPLICANT:  Stiegler, Gary
   APPLICANT:  Gaines, Patrick J.
   APPLICANT:  Silver, Gary
   TITLE OF INVENTION:  FLEA AMINOPEPTIDASE PROTEINS AND USES THEREOF
   NUMBER OF SEQUENCES:  190
   CORRESPONDENCE ADDRESS:
     ADDRESSEE:  Sheridan Ross P.C.
     STREET:  1700 Lincoln Street, Suite 3500
     CITY:  Denver
     STATE:  Colorado
     COUNTRY:  USA
     ZIP:  80203
   COMPUTER READABLE FORM:
     MEDIUM TYPE:  Floppy disk
     COMPUTER:  IBM PC compatible
     OPERATING SYSTEM:  PC-DOS/MS-DOS
     SOFTWARE:  PatentIn Release #1.0, Version #1.30
   CURRENT APPLICATION DATA:
     APPLICATION NUMBER:  US/08/906,616
     FILING DATE:  05-AUG-1997
     CLASSIFICATION:  536
   ATTORNEY/AGENT INFORMATION:
     NAME:  Connell, Gary J.
     REGISTRATION NUMBER:  32,020
     REFERENCE/DOCKET NUMBER:  2618-25-C2-3
   TELECOMMUNICATION INFORMATION:
     TELEPHONE:  (303) 863-9700
     TELEFAX:  (303) 863-0223
  INFORMATION FOR SEQ ID NO:  128:
   SEQUENCE CHARACTERISTICS:
     LENGTH:  815 base pairs
     TYPE:  nucleic acid
     STRANDEDNESS:  single
     TOPOLOGY:  linear
   MOLECULE TYPE:  cDNA
   FEATURE:
     NAME/KEY:  CDS
     LOCATION:  1..762
     OTHER INFORMATION:  /note= "At pos. bp 453, change A to
     OTHER INFORMATION:  M; at 454, change G to V; at 456, G to V; at 457, A to M; at
     OTHER INFORMATION:  460, A to R; at 470, G to S; at 493, A to R.  At pos. aa 120,
     OTHER INFORMATION:  136, 152, 153, 154, 157 and 165, substitute Xaa."
US-08-906-616-128


  Query Match             0.5%;  Score 16;  DB 3;  Length 815;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1478 aagaagtcagtaccca 1493
          ||||||||||||||||
Db    181 AAGAAGTCAGTACCCA 166


RESULT  68
US-08-639-075A-128/c
 Sequence 128, Application US/08639075A
 Patent No. 6150125
  GENERAL INFORMATION:
    APPLICANT:  Grieve, Robert B.
    APPLICANT:  Rushlow, Keith E.
    APPLICANT:  Wu Hunter, Shirley
    APPLICANT:  Frank, Glenn R.
    APPLICANT:  Stiegler, Gary
    APPLICANT:  Gaines, Patrick J.
    APPLICANT:  Silver, Gary
    TITLE OF INVENTION:  FLEA PROTEASE PROTEINS, NUCLEIC ACID
;    TITLE OF INVENTION:  MOLECULES AND USES THEREOF
;    NUMBER OF SEQUENCES:  190
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sheridan Ross & McIntosh
;      STREET:  1700 Lincoln Street, Suite 3500
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80203
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/639,075A
;      FILING DATE:  24-APR-1996
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Connell, Gary J.
;      REGISTRATION NUMBER:  32,020
;      REFERENCE/DOCKET NUMBER:  2618-25-C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 863-9700
;      TELEFAX:  (303) 863-0223
;  INFORMATION FOR SEQ ID NO:  128:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  815 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..762
;      OTHER INFORMATION:  /note= "At pos. bp 453, change A to
;      OTHER INFORMATION:  M; at 454, change G to V; at 456, G to V; at 457, A to M;
;      OTHER INFORMATION:  460, A to R; at 470, G to S; at 493, A to R.  At pos. aa 1
;      OTHER INFORMATION:  136, 152, 153, 154, 157 and 165, substitute Xaa."
US-08-639-075A-128


  Query Match             0.5%;  Score 16;  DB 3;  Length 815;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1478 aagaagtcagtaccca 1493
          ||||||||||||||||
Db    181 AAGAAGTCAGTACCCA 166


RESULT  69
US-08-651-136C-19
; Sequence 19, Application US/08651136C
; Patent No. 6001639
;  GENERAL INFORMATION:
;    APPLICANT:  Schulein, Martin
;    APPLICANT:  Andersen, Lene N.
;    APPLICANT:  Lassen, Soren F.
;    APPLICANT:  Kauppinen, Markus S.
;    APPLICANT:  Lange, Lene
;    APPLICANT:  Nielsen, Ruby I.
;    APPLICANT:  Ihara, Michiko
;    APPLICANT:  Takagi, Shinobu
;    TITLE OF INVENTION:  No. 6001639el Endoglucanases
;    NUMBER OF SEQUENCES:  109
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6001639o No. 6001639disk of No. 6001639th America, Inc.
;      STREET:  405 Lexington Avenue, 64th Floor
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  United States of America
;      ZIP:  10174-6401
```

```
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/651,136C
      FILING DATE:  21-MAY-1996
      CLASSIFICATION:  435
    ATTORNEY/AGENT INFORMATION:
      NAME:  Lambiris, Elias J.
      REGISTRATION NUMBER:  33,728
      REFERENCE/DOCKET NUMBER:  4366.200-US
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  212-867-0123
      TELEFAX:  212-878-9655
  INFORMATION FOR SEQ ID NO:  19:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  1031 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    MOLECULE TYPE:  cDNA
    FEATURE:
      NAME/KEY:  CDS
      LOCATION:  11..889
S-08-651-136C-19


  Query Match             0.5%;  Score 16;  DB 3;  Length 1031;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

y   1409 tcccaacttccagcag 1424
         ||||||||||||||||
b    536 TCCCAACTTCCAGCAG 551


ESULT  70
S-08-651-136C-17
 Sequence 17, Application US/08651136C
 Patent No. 6001639
  GENERAL INFORMATION:
    APPLICANT:  Schulein, Martin
    APPLICANT:  Andersen, Lene N.
    APPLICANT:  Lassen, Soren F.
    APPLICANT:  Kauppinen, Markus S.
    APPLICANT:  Lange, Lene
    APPLICANT:  Nielsen, Ruby I.
    APPLICANT:  Ihara, Michiko
    APPLICANT:  Takagi, Shinobu
    TITLE OF INVENTION:  No. 6001639el Endoglucanases
    NUMBER OF SEQUENCES:  109
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  No. 6001639o No. 6001639disk of No. 6001639th America, Inc.
      STREET:  405 Lexington Avenue, 64th Floor
      CITY:  New York
      STATE:  New York
      COUNTRY:  United States of America
      ZIP:  10174-6401
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/651,136C
      FILING DATE:  21-MAY-1996
      CLASSIFICATION:  435
    ATTORNEY/AGENT INFORMATION:
      NAME:  Lambiris, Elias J.
      REGISTRATION NUMBER:  33,728
      REFERENCE/DOCKET NUMBER:  4366.200-US
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  212-867-0123
      TELEFAX:  212-878-9655
  INFORMATION FOR SEQ ID NO:  17:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  1048 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    MOLECULE TYPE:  cDNA
    FEATURE:
      NAME/KEY:  CDS
      LOCATION:  13..906
US-08-651-136C-17


  Query Match             0.5%;  Score 16;  DB 3;  Length 1048;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1409 tcccaacttccagcag 1424
          ||||||||||||||||
Db    553 TCCCAACTTCCAGCAG 568


RESULT  71
US-08-933-750C-81/c
 Sequence 81, Application US/08933750C
 Patent No. 5932442
  GENERAL INFORMATION:
    APPLICANT:  Lal, Preeti
    APPLICANT:  Hillman, Jennifer L.
    APPLICANT:  Bandman, Olga
    APPLICANT:  Shah, Purvi
    APPLICANT:  Au-Young, Janice
    APPLICANT:  Yue, Henry
    APPLICANT:  Guegler, Karl J.
    APPLICANT:  Corley, Neil C.
    TITLE OF INVENTION:  HUMAN REGULATORY MOLECULES
    NUMBER OF SEQUENCES:  98
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
      STREET:  3174 Porter Drive
      CITY:  Palo Alto
      STATE:  CA
      COUNTRY:  USA
      ZIP:  94304
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Diskette
      COMPUTER:  IBM Compatible
      OPERATING SYSTEM:  DOS
      SOFTWARE:  FastSEQ for Windows Version 2.0
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/933,750C
      FILING DATE:  September 23, 1997
      CLASSIFICATION:  536
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:
      FILING DATE:
    ATTORNEY/AGENT INFORMATION:
      NAME:  Billings, Lucy J.
      REGISTRATION NUMBER:  36,749
      REFERENCE/DOCKET NUMBER:  PF-0356 US
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  415-855-0555
      TELEFAX:  415-845-4166
      TELEX:
  INFORMATION FOR SEQ ID NO:  81:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  1152 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
```

```
    IMMEDIATE SOURCE:
      LIBRARY:  BLADNOT03
      CLONE:  1602473
US-08-933-750C-81


  Query Match             0.5%;  Score 16;  DB 2;  Length 1152;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2168 ccaagccatcagcgtg 2183
          ||||||||||||||||
Db    928 CCAAGCCATCAGCGTG 913


RESULT  72
US-09-234-613-81/c
; Sequence 81, Application US/09234613
; Patent No. 6132973
;  GENERAL INFORMATION:
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Au-Young, Janice
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  HUMAN REGULATORY MOLECULES
;    NUMBER OF SEQUENCES:  98
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/234,613
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/933,750
;      FILING DATE:  September 23, 1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0356 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  81:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1152 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  BLADNOT03
;      CLONE:  1602473
US-09-234-613-81


  Query Match             0.5%;  Score 16;  DB 3;  Length 1152;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2168 ccaagccatcagcgtg 2183
          ||||||||||||||||
Db    928 CCAAGCCATCAGCGTG 913


RESULT  73
US-08-739-485-4
; Sequence 4, Application US/08739485
; Patent No. 5863898
;  GENERAL INFORMATION:
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Bandman, Olga
;    TITLE OF INVENTION:  NOVEL HUMAN LIM PROTEINS
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/739,485
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0142 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1225 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  Consensus
;      CLONE:  Consensus
US-08-739-485-4


  Query Match             0.5%;  Score 16;  DB 2;  Length 1225;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1835 ccaggaggtcctgcac 1850
          ||||||||||||||||
Db    441 CCAGGAGGTCCTGCAC 456


RESULT  74
US-08-360-758-1/c
; Sequence 1, Application US/08360758
; Patent No. 6074863
;  GENERAL INFORMATION:
;    APPLICANT:  Svendsen, Allan
;    APPLICANT:  Pathar, Shamkant A
;    APPLICANT:  Egel-Mitani, Michi
;    APPLICANT:  Borch, Kim
;    APPLICANT:  Clausen, Ib G
```

```
    APPLICANT:  Hansen, Mogens T
    TITLE OF INVENTION:  C. ANTARCTICA LIPASE AND LIPASE VARIANTS
    NUMBER OF SEQUENCES:  11
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  No. 6074863o No. 6074863disk of No. 6074863th America, Inc.
      STREET:  405 Lexington Avenue, 64th Floor
      CITY:  New York
      STATE:  New York
      COUNTRY:  United States of America
      ZIP:  10174-6401
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Tape
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/360,758
      FILING DATE:  22-DEC-1994
      CLASSIFICATION:  435
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  DK PCT/DK93/00225
      FILING DATE:  03-JUN-1993
    ATTORNEY/AGENT INFORMATION:
      NAME:  Lambiris, Elias J.
      REGISTRATION NUMBER:  33,728
      REFERENCE/DOCKET NUMBER:  3748.204-US
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  212-867-0123
      TELEFAX:  212-878-9655
  INFORMATION FOR SEQ ID NO:  1:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  1329 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    MOLECULE TYPE:  DNA (genomic)
US-08-360-758-1


  Query Match             0.5%;  Score 16;  DB 3;  Length 1329;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1143 acccagcacttggtcc 1158
           ||||||||||||||||
Db    1245 ACCCAGCACTTGGTCC 1230


RESULT  75
US-08-889-425-3/c
  Sequence 3, Application US/08889425
  Patent No. 6153403
   GENERAL INFORMATION:
     APPLICANT:  Lim, Bing
     APPLICANT:  Adra, Chaker N.
     TITLE OF INVENTION:  A Lysosomal-Associated Multispanning
     TITLE OF INVENTION:  Membrane Protein, LAPTM5 and a Nucleic Acid Encoding
     TITLE OF INVENTION:  LAPTM5
     NUMBER OF SEQUENCES:  6
     CORRESPONDENCE ADDRESS:
       ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
       STREET:  Two Militia Drive
       CITY:  Lexington
       STATE:  Massachusetts
       COUNTRY:  USA
       ZIP:  02173
     COMPUTER READABLE FORM:
       MEDIUM TYPE:  Floppy disk
       COMPUTER:  IBM PC compatible
       OPERATING SYSTEM:  PC-DOS/MS-DOS
       SOFTWARE:  PatentIn Release #1.0, Version #1.30
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER:  US/08/889,425
       FILING DATE:  08-JUL-1997
       CLASSIFICATION:  530
     ATTORNEY/AGENT INFORMATION:
       NAME:  Brook, David E.
       REGISTRATION NUMBER:  22,592
       REFERENCE/DOCKET NUMBER:  BIH96-09pA
     TELECOMMUNICATION INFORMATION:
       TELEPHONE:  (781) 861-6240
       TELEFAX:  (781) 861-9540
   INFORMATION FOR SEQ ID NO:  3:
     SEQUENCE CHARACTERISTICS:
       LENGTH:  1333 base pairs
       TYPE:  nucleic acid
       STRANDEDNESS:  double
       TOPOLOGY:  linear
     MOLECULE TYPE:  cDNA
     FEATURE:
       NAME/KEY:  CDS
       LOCATION:  58..840
US-08-889-425-3


  Query Match             0.5%;  Score 16;  DB 3;  Length 1333;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    534 gaactggctgtgcggc 549
          ||||||||||||||||
Db    606 GAACTGGCTGTGCGGC 591


RESULT  76
US-09-032-372-9/c
  Sequence 9, Application US/09032372
  Patent No. 6008337
   GENERAL INFORMATION:
     APPLICANT:  Bandman, Olga
     APPLICANT:  Hillman, Jennifer L.
     APPLICANT:  Corley, Neil C.
     APPLICANT:  Guegler, Karl J.
     APPLICANT:  Yue, Henry
     APPLICANT:  Lal, Preeti
     TITLE OF INVENTION:  CELL CYCLE RELATED PROTEINS
     NUMBER OF SEQUENCES:  13
     CORRESPONDENCE ADDRESS:
       ADDRESSEE:  Incyte Pharmaceuticals, Inc.
       STREET:  3174 Porter Drive
       CITY:  Palo Alto
       STATE:  CA
       COUNTRY:  USA
       ZIP:  94304
     COMPUTER READABLE FORM:
       MEDIUM TYPE:  Diskette
       COMPUTER:  IBM Compatible
       OPERATING SYSTEM:  DOS
       SOFTWARE:  FastSEQ for Windows Version 2.0
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER:  US/09/032,372
       FILING DATE:  Herewith
       CLASSIFICATION:
     PRIOR APPLICATION DATA:
       APPLICATION NUMBER:
       FILING DATE:
     ATTORNEY/AGENT INFORMATION:
       NAME:  Billings, Lucy J
       REGISTRATION NUMBER:  36,749
       REFERENCE/DOCKET NUMBER:  PF-0478 US
     TELECOMMUNICATION INFORMATION:
       TELEPHONE:  650-855-0555
       TELEFAX:  650-845-4166
       TELEX:
   INFORMATION FOR SEQ ID NO:  9:
     SEQUENCE CHARACTERISTICS:
```

```
      LENGTH:  1341 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    IMMEDIATE SOURCE:
      LIBRARY:  BRAITUT21
      CLONE:  2522306
S-09-032-372-9


  Query Match             0.5%;  Score 16;  DB 3;  Length 1341;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2624 cacggtcccccaggag 2639
          ||||||||||||||||
Db    794 CACGGTCCCCCAGGAG 779


RESULT  77
S-08-374-155A-7
 Sequence 7, Application US/08374155A
 Patent No. 5786140
  GENERAL INFORMATION:
    APPLICANT:  Mattes, Ralf
    APPLICANT:  Klein, Kathrin
    APPLICANT:  Schiweck, Hubert
    APPLICANT:  Kunz, Markwart
    APPLICANT:  Munir, Mohammed
    TITLE OF INVENTION:  Preparation of Acariogenic Sugar
    TITLE OF INVENTION:  Substitutes
    NUMBER OF SEQUENCES:  26
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
      ADDRESSEE:  Dunner
      STREET:  1300 I Street, N.W.
      CITY:  Washington
      STATE:  D.C.
      COUNTRY:  USA
      ZIP:  20005-3315
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/374,155A
      FILING DATE:  18-JAN-1995
      CLASSIFICATION:  435
    ATTORNEY/AGENT INFORMATION:
      NAME:  Forman, David S
      REGISTRATION NUMBER:  33,694
      REFERENCE/DOCKET NUMBER:  05638.0006-00000
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (202) 408-4000
      TELEFAX:  (202) 408-4400
  INFORMATION FOR SEQ ID NO:  7:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  1362 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    MOLECULE TYPE:  DNA (genomic)
S-08-374-155A-7


  Query Match             0.5%;  Score 16;  DB 1;  Length 1362;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2213 gctgaaccacttcagc 2228
          ||||||||||||||||
Db    594 GCTGAACCACTTCAGC 609


RESULT  78
US-08-785-396-7
; Sequence 7, Application US/08785396
; Patent No. 5985622
;  GENERAL INFORMATION:
;    APPLICANT:  Mattes, Ralf
;    APPLICANT:  Klein, Kathrin
;    APPLICANT:  Schiweck, Hubert
;    APPLICANT:  Kunz, Markwart
;    APPLICANT:  Munir, Mohammed
;    TITLE OF INVENTION:  Preparation of Acariogenic Sugar
;    TITLE OF INVENTION:  Substitutes
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/785,396
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/374,155
;      FILING DATE:  18-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Forman, David S
;      REGISTRATION NUMBER:  33,694
;      REFERENCE/DOCKET NUMBER:  05638.0006-00000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 408-4000
;      TELEFAX:  (202) 408-4400
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1362 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-785-396-7


  Query Match             0.5%;  Score 16;  DB 2;  Length 1362;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2213 gctgaaccacttcagc 2228
          ||||||||||||||||
Db    594 GCTGAACCACTTCAGC 609


RESULT  79
US-08-458-023B-1/c
; Sequence 1, Application US/08458023B
; Patent No. 5667990
;  GENERAL INFORMATION:
;    APPLICANT:  Berka, Randy M.
;    APPLICANT:  Yoder, Wendy
;    APPLICANT:  Takagi, Shinobu
;    APPLICANT:  Boominathan, Karuppan C.
;    TITLE OF INVENTION:  ASPERGILLUS EXPRESSION SYSTEM
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
```

```
      ADDRESSEE:  No. 5667990o No. 5667990disk of No. 5667990th America, Inc.
      STREET:  405 Lexington Avenue
      CITY:  New York
      STATE:  New York
      COUNTRY:  USA
      ZIP:  10174-6201
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/458,023B
      FILING DATE:  01-JUN-1995
      CLASSIFICATION:  435
    ATTORNEY/AGENT INFORMATION:
      NAME:  Lowney Dr., Karen A.
      REGISTRATION NUMBER:  31,274
      REFERENCE/DOCKET NUMBER:  4086.010-US
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  212-867-0123
      TELEFAX:  212-878-9655
  INFORMATION FOR SEQ ID NO:  1:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  1389 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    MOLECULE TYPE:  DNA (genomic)
    HYPOTHETICAL:  NO
    ANTI-SENSE:  NO
    ORIGINAL SOURCE:
      ORGANISM:  Candida antarctica
      INDIVIDUAL ISOLATE:  DSM 3855
    FEATURE:
      NAME/KEY:  CDS
      LOCATION:  1..1389
S-08-458-023B-1


  Query Match             0.5%;  Score 16;  DB 1;  Length 1389;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1143 acccagcacttggtcc 1158
          ||||||||||||||||
Db   1305 ACCCAGCACTTGGTCC 1290


RESULT  80
US-09-111-556A-1/c
  Sequence 1, Application US/09111556A
  Patent No. 6020180
  GENERAL INFORMATION:
    APPLICANT:  Svendsen, Allan
    APPLICANT:  Pathar, Shamkant A
    APPLICANT:  Egel-Mitani, Michi
    APPLICANT:  Borch, Kim
    APPLICANT:  Clausen, Ib G
    APPLICANT:  Hansen, Mogens T
    TITLE OF INVENTION:  C. ANTARCTICA LIPASE AND LIPASE VARIANTS
    NUMBER OF SEQUENCES:  11
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  No. 6020180o No. 6020180disk of No. 6020180th America, Inc.
      STREET:  405 Lexington Avenue, 64th Floor
      CITY:  New York
      STATE:  New York
      COUNTRY:  United States of America
      ZIP:  10174-6401
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Tape
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/09/111,556A
      FILING DATE:  22-DEC-1994
      CLASSIFICATION:
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  DK PCT/DK93/00225
      FILING DATE:  03-JUN-1993
    ATTORNEY/AGENT INFORMATION:
      NAME:  Lambiris, Elias J.
      REGISTRATION NUMBER:  33,728
      REFERENCE/DOCKET NUMBER:  3748.214-US
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  212-867-0123
      TELEFAX:  212-878-9655
  INFORMATION FOR SEQ ID NO:  1:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  1389 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    MOLECULE TYPE:  DNA (genomic)
US-09-111-556A-1


  Query Match             0.5%;  Score 16;  DB 3;  Length 1389;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1143 acccagcacttggtcc 1158
          ||||||||||||||||
Db   1305 ACCCAGCACTTGGTCC 1290


RESULT  81
US-08-135-510-4/c
; Sequence 4, Application US/08135510
; Patent No. 5420028
;  GENERAL INFORMATION:
;    APPLICANT:  CHIANG, John Young Ling
;    TITLE OF INVENTION:  Truncated Human Cholesterol
;    TITLE OF INVENTION:  7a-Hydroxylase, Method of Production and Use Thereof
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington, D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/135,510
;      FILING DATE:  13-OCT-1993
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SANDERCOCK, Colin G.
;      REGISTRATION NUMBER:  31,298
;      REFERENCE/DOCKET NUMBER:  18748/176 HOCE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1524 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-135-510-4
```

```
  Query Match             0.5%;  Score 16;  DB 1;  Length 1524;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    499 caatcaaaatattttc 514
          ||||||||||||||||
Db    252 CAATCAAAATATTTTC 237


RESULT  82
US-08-483-852-4/c
; Sequence 4, Application US/08483852
; Patent No. 5650286
;  GENERAL INFORMATION:
;    APPLICANT:  CHIANG, John Young Ling
;    TITLE OF INVENTION:  Genomic DNA of Human Cholesterol
;    TITLE OF INVENTION:  7a-Hydroxylase and Methods of Using It
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,852
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/361,458
;      FILING DATE:  21-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/135,511
;      FILING DATE:  13-OCT-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/135,488
;      FILING DATE:  13-OCT-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/135,510
;      FILING DATE:  13-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SANDERCOCK, Colin G.
;      REGISTRATION NUMBER:  31,298
;      REFERENCE/DOCKET NUMBER:  18748/220 HOCE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1524 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-483-852-4


  Query Match             0.5%;  Score 16;  DB 1;  Length 1524;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    499 caatcaaaatattttc 514
          ||||||||||||||||
Db    252 CAATCAAAATATTTTC 237


RESULT  83
US-08-477-953-4/c
; Sequence 4, Application US/08477953
; Patent No. 5677159
;  GENERAL INFORMATION:
;    APPLICANT:  CHIANG, John Young Ling
;    TITLE OF INVENTION:  Genomic DNA of Human Cholesterol
;    TITLE OF INVENTION:  7a-Hydroxylase and Methods of Using It
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,953
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/361,458
;      FILING DATE:  21-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/135,511
;      FILING DATE:  13-OCT-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/135,488
;      FILING DATE:  13-OCT-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/135,510
;      FILING DATE:  13-OCT-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SANDERCOCK, Colin G.
;      REGISTRATION NUMBER:  31,298
;      REFERENCE/DOCKET NUMBER:  18748/219 HOCE
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1524 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-477-953-4


  Query Match             0.5%;  Score 16;  DB 1;  Length 1524;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    499 caatcaaaatattttc 514
          ||||||||||||||||
Db    252 CAATCAAAATATTTTC 237


RESULT  84
US-08-409-122-1
; Sequence 1, Application US/08409122
; Patent No. 5820870
;  GENERAL INFORMATION:
;    APPLICANT:  JOYCE, JAMES G.
;    APPLICANT:  GEORGE, HUGH A.
;    APPLICANT:  HOFMANN, KATHRYN J.
;    APPLICANT:  JANSEN, KATHRIN U.
;    APPLICANT:  NEEPER, MICHAEL P.
;    TITLE OF INVENTION:  DNA ENCODING HUMAN PAPILLOMAVIRUS TYPE 18 VACCINE
```

```
    NUMBER OF SEQUENCES:  16
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  CHRISTINE E. CARTY - MERCK & CO., INC.
      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
      CITY:  RAHWAY
      STATE:  NJ
      COUNTRY:  US
      ZIP:  07065-0907
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Diskette
      COMPUTER:  IBM Compatible
      OPERATING SYSTEM:  DOS
      SOFTWARE:  FastSEQ Version 1.5
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/409,122
      FILING DATE:
      CLASSIFICATION:  435
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  08/408,669
      FILING DATE:  22-MAR-1995
    ATTORNEY/AGENT INFORMATION:
      NAME:  CARTY, CHRISTINE E
      REGISTRATION NUMBER:  36,099
      REFERENCE/DOCKET NUMBER:  19425
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  908-594-6734
      TELEFAX:  908-594-4720
      TELEX:
  INFORMATION FOR SEQ ID NO:  1:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  1524 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    MOLECULE TYPE:  cDNA
    HYPOTHETICAL:  NO
    ANTI-SENSE:  NO
    FRAGMENT TYPE:
    ORIGINAL SOURCE:
R-08-409-122-1


  Query Match             0.5%;  Score 16;  DB 1;  Length 1524;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

y    227 cctgcaggtggtggca 242
        ||||||||||||||||
b    157 CCTGCAGGTGGTGGCA 172


SULT  85
S-08-408-669-1
 Sequence 1, Application US/08408669
 Patent No. 5840306
  GENERAL INFORMATION:
    APPLICANT:  HOFMANN, KATHRYN J.
    APPLICANT:  JANSEN, KATHRIN U.
    APPLICANT:  NEEPER, MICHAEL P.
    TITLE OF INVENTION:  DNA ENCODING HUMAN PAPILLOMAVIRUS TYPE 18
    NUMBER OF SEQUENCES:  16
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  CHRISTINE E. CARTY - MERCK & CO., INC.
      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
      CITY:  RAHWAY
      STATE:  NJ
      COUNTRY:  US
      ZIP:  07065-0907
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Diskette
      COMPUTER:  IBM Compatible
      OPERATING SYSTEM:  DOS
      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/408,669
;      FILING DATE:  22-MAR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CARTY, CHRISTINE E
;      REGISTRATION NUMBER:  36,099
;      REFERENCE/DOCKET NUMBER:  19424
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-594-6734
;      TELEFAX:  908-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1524 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
US-08-408-669-1


  Query Match             0.5%;  Score 16;  DB 2;  Length 1524;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    227 cctgcaggtggtggca 242
        ||||||||||||||||
Db    157 CCTGCAGGTGGTGGCA 172


RESULT  86
US-08-477-952-4/c
; Sequence 4, Application US/08477952
; Patent No. 5851780
;  GENERAL INFORMATION:
;    APPLICANT:  CHIANG, John Young Ling
;    TITLE OF INVENTION:  Genomic DNA of Human Cholesterol
;    TITLE OF INVENTION:  7a-Hydroxylase and Methods of Using It
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,952
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/361,458
;      FILING DATE:  21-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/135,511
;      FILING DATE:  13-OCT-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/135,488
;      FILING DATE:  13-OCT-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/135,510
```

```
      FILING DATE:  13-OCT-1993
    ATTORNEY/AGENT INFORMATION:
      NAME:  SANDERCOCK, Colin G.
      REGISTRATION NUMBER:  31,298
      REFERENCE/DOCKET NUMBER:  18748/221 HOCE
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (202)672-5300
      TELEFAX:  (202)672-5399
      TELEX:  904136
  INFORMATION FOR SEQ ID NO:  4:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  1524 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  double
      TOPOLOGY:  linear
S-08-477-952-4


  Query Match             0.5%;  Score 16;  DB 2;  Length 1524;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    499 caatcaaaatattttc 514
          ||||||||||||||||
Db    252 CAATCAAAATATTTTC 237


RESULT  87
S-08-386-727-5/c
 Sequence 5, Application US/08386727
 Patent No. 5792647
  GENERAL INFORMATION:
    APPLICANT:  ROSEMAN, SAUL
    APPLICANT:  BASSLER, BONNIE
    APPLICANT:  KEYHANI, NEMAT O.
    APPLICANT:  CHITLARU, EDITH
    APPLICANT:  ROWE, CHRIS
    APPLICANT:  YU, CHARLES
    TITLE OF INVENTION:  BACTERIAL CATABOLISM OF CHITIN
    NUMBER OF SEQUENCES:  8
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  CUSHMAN, DARBY & CUSHMAN
      STREET:  1100 NEW YORK AVENUE, N.W.
      CITY:  WASHINGTON
      STATE:  D.C.
      COUNTRY:  USA
      ZIP:  20005
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.25
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/386,727
      FILING DATE:
      CLASSIFICATION:  435
    ATTORNEY/AGENT INFORMATION:
      NAME:  HOBBS, ANN S.
      REGISTRATION NUMBER:  36,830
      REFERENCE/DOCKET NUMBER:  4130/206916
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  202-861-3000
      TELEFAX:  202-822-0944
      TELEX:  6714627 CUSH
  INFORMATION FOR SEQ ID NO:  5:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  1713 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  single
      TOPOLOGY:  linear
    MOLECULE TYPE:  DNA (genomic)
S-08-386-727-5


  Query Match             0.5%;  Score 16;  DB 1;  Length 1713;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1176 tcagttcacaaccttc 1191
          ||||||||||||||||
Db    241 TCAGTTCACAACCTTC 226


RESULT  88
US-08-600-452A-5/c
; Sequence 5, Application US/08600452A
; Patent No. 5985644
;  GENERAL INFORMATION:
;    APPLICANT:  ROSEMAN, SAUL
;    APPLICANT:  BASSLER, BONNIE
;    APPLICANT:  KEYHANI, NEMAT O.
;    APPLICANT:  CHITLARU, EDITH
;    APPLICANT:  ROWE, CHRIS
;    APPLICANT:  YU, CHARLES
;    TITLE OF INVENTION:  BACTERIAL CATABOLISM OF CHITIN
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FISH & RICHARDSON P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/600,452A
;      FILING DATE:  13-FEB-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haile, Lisa A.
;      REGISTRATION NUMBER:  38,347
;      REFERENCE/DOCKET NUMBER:  07662/005001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 678-5070
;      TELEFAX:  (619) 678-5099
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1713 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-600-452A-5


  Query Match             0.5%;  Score 16;  DB 2;  Length 1713;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1176 tcagttcacaaccttc 1191
          ||||||||||||||||
Db    241 TCAGTTCACAACCTTC 226


RESULT  89
US-08-467-948A-2/c
; Sequence 1, Application US/08467948A
; Patent No. 5998164
;  GENERAL INFORMATION:
;    APPLICANT:  LI, YI
;    APPLICANT:  CAO, LIANG
```

```
    APPLICANT:  NI, JIAN
    APPLICANT:  GENTZ, REINER
    APPLICANT:  BULT, CAROL J.
    APPLICANT:  SUTTON III, GRANGER G.
    APPLICANT:  ROSEN, CRAIG A.
    TITLE OF INVENTION:  Polynucleotides Encoding Human G-Protein
    TITLE OF INVENTION:  Coupled Receptor GPR2
    NUMBER OF SEQUENCES:  30
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
      STREET:  1100 NEW YORK AVE., NW, SUITE 600
      CITY:  WASHINGTON
      STATE:  DC
      COUNTRY:  USA
      ZIP:  20005
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  FLOPPY DISK
      COMPUTER:  IBM PC COMPATIBLE
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PATENTIN RELEASE #1.0, VERSION #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/467,948A
      FILING DATE:  06-JUN-1995
      CLASSIFICATION:  435
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  PCT/US95/04079
      FILING DATE:  30-MAR-1995
    ATTORNEY/AGENT INFORMATION:
      NAME:  STEFFE, ERIC K.
      REGISTRATION NUMBER:  36,688
      REFERENCE/DOCKET NUMBER:  1488.1140003/EKS/KLM
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  202-371-2600
      TELEFAX:  202-371-2540
  INFORMATION FOR SEQ ID NO:  1:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  1713 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  both
      TOPOLOGY:  both
    MOLECULE TYPE:  cDNA
    FEATURE:
      NAME/KEY:  CDS
      LOCATION:  116..1003
US-08-467-948A-1


  Query Match             0.5%;  Score 16;  DB 2;  Length 1713;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    824 agcaaaggagatgggc 839
          ||||||||||||||||
Db    409 AGCAAAGGAGATGGGC 394


RESULT  90
US-08-467-947A-1/c
  Sequence 1, Application US/08467947A
  Patent No. 6090575
  GENERAL INFORMATION:
    APPLICANT:  LI, YI
    APPLICANT:  CAO, LIANG
    APPLICANT:  NI, JIAN
    APPLICANT:  GENTZ, REINER
    APPLICANT:  BULT, CAROL J.
    APPLICANT:  SUTTON III, GRANGER G.
    APPLICANT:  ROSEN, CRAIG A.
    TITLE OF INVENTION:  Polynucleotides Encoding Human G-Protein
    TITLE OF INVENTION:  Coupled Receptor GPR1
    NUMBER OF SEQUENCES:  30
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
      STREET:  1100 NEW YORK AVE., NW, SUITE 600
      CITY:  WASHINGTON
      STATE:  DC
      COUNTRY:  USA
      ZIP:  20005
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  FLOPPY DISK
      COMPUTER:  IBM PC COMPATIBLE
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PATENTIN RELEASE #1.0, VERSION #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/467,947A
      FILING DATE:  06-JUN-1995
      CLASSIFICATION:  435
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  PCT/US95/04079
      FILING DATE:  30-MAR-1995
    ATTORNEY/AGENT INFORMATION:
      NAME:  STEFFE, ERIC K.
      REGISTRATION NUMBER:  36,688
      REFERENCE/DOCKET NUMBER:  1488.1140002/EKS/KLM
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  202-371-2600
      TELEFAX:  202-371-2540
  INFORMATION FOR SEQ ID NO:  1:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  1713 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  both
      TOPOLOGY:  both
    MOLECULE TYPE:  cDNA
    FEATURE:
      NAME/KEY:  CDS
      LOCATION:  116..1003
US-08-467-947A-1


  Query Match             0.5%;  Score 16;  DB 3;  Length 1713;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    824 agcaaaggagatgggc 839
          ||||||||||||||||
Db    409 AGCAAAGGAGATGGGC 394


RESULT  91
US-07-688-352C-13
; Sequence 13, Application US/07688352C
; Patent No. 5527896
; GENERAL INFORMATION:
;   APPLICANT:  Wigler, Michael H.
;   APPLICANT:  Colicelli, John J.
;   TITLE OF INVENTION:  Cloning by Complementation and Related
;   TITLE OF INVENTION:  Processes
;   NUMBER OF SEQUENCES:  57
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;     ADDRESSEE:  Bicknell
;     STREET:  Two First National Plaza, 20 South Clark
;     STREET:  Street
;     CITY:  Chicago
;     STATE:  Illinois
;     COUNTRY:  USA
;     ZIP:  60603
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/07/688,352C
;     FILING DATE:  19910419
```

```
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/511,715
;      FILING DATE:  20-APR-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25447
;      REFERENCE/DOCKET NUMBER:  27805/30197
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 346-5750
;      TELEFAX:  (312) 984-9740
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1721 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  60..1274
US-07-688-352C-13


  Query Match             0.5%;  Score 16;  DB 1;  Length 1721;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2731 cacagccgcgggccag 2746
          ||||||||||||||||
Db    266 CACAGCCGCGGGCCAG 281


RESULT  92
US-08-474-379C-13
; Sequence 13, Application US/08474379C
; Patent No. 5977305
;  GENERAL INFORMATION:
;    APPLICANT:  Wigler, Michael H.
;    APPLICANT:  Colicelli, John J.
;    TITLE OF INVENTION:  CLONING BY COMPLEMENTATION AND RELATED
;    TITLE OF INVENTION:  PROCESSES
;    NUMBER OF SEQUENCES:  88
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive/6300 Sears Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/474,379C
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/511,715
;      FILING DATE:  20-APR-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/206,188
;      FILING DATE:  01-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/688,352
;      FILING DATE:  19-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clough, David W.
;      REGISTRATION NUMBER:  36,107
;      REFERENCE/DOCKET NUMBER:  27866/32771
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1721 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  66..1274
US-08-474-379C-13


  Query Match             0.5%;  Score 16;  DB 2;  Length 1721;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2731 cacagccgcgggccag 2746
          ||||||||||||||||
Db    266 CACAGCCGCGGGCCAG 281


RESULT  93
US-09-146-249A-13
; Sequence 13, Application US/09146249A
; Patent No. 6069240
;  GENERAL INFORMATION:
;    APPLICANT:  Wigler, Michael H.
;    APPLICANT:  Colicelli, John J.
;    TITLE OF INVENTION:  Cloning by Complementation and Related
;    TITLE OF INVENTION:  Processes
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/146,249A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/511,715
;      FILING DATE:  20-APR-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clough, David W.
;      REGISTRATION NUMBER:  36,107
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1721 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  66..1274
US-09-146-249A-13
```

```
  Query Match             0.5%;  Score 16;  DB 3;  Length 1721;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2731 cacagccgcgggccag 2746
          ||||||||||||||||
Db    266 CACAGCCGCGGGCCAG 281


RESULT  94
US-08-206-188B-13
  Sequence 13, Application US/08206188B
  Patent No. 6100025
   GENERAL INFORMATION:
     APPLICANT:  Wigler, Michael H.
     APPLICANT:  Colicelli, John J.
     TITLE OF INVENTION:  Cloning by Complementation and Related
     TITLE OF INVENTION:  Processes
     NUMBER OF SEQUENCES:  84
     CORRESPONDENCE ADDRESS:
       ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
       STREET:  6300 Sears Tower, 233 South Wacker Drive
       CITY:  Chicago
       STATE:  Illinois
       COUNTRY:  United States of America
       ZIP:  60606-6402
     COMPUTER READABLE FORM:
       MEDIUM TYPE:  Floppy disk
       COMPUTER:  IBM PC compatible
       OPERATING SYSTEM:  PC-DOS/MS-DOS
       SOFTWARE:  PatentIn Release #1.0, Version #1.25
     CURRENT APPLICATION DATA:
       APPLICATION NUMBER:  US/08/206,188B
       FILING DATE:  01-MAR-1994
       CLASSIFICATION:  435
     PRIOR APPLICATION DATA:
       APPLICATION NUMBER:  US 07/511,715
       FILING DATE:  20-APR-1990
     ATTORNEY/AGENT INFORMATION:
       NAME:  Clough, David W.
       REGISTRATION NUMBER:  36107
     TELECOMMUNICATION INFORMATION:
       TELEPHONE:  312/474-6300
       TELEFAX:  312-474-0448
       TELEX:  25-3856
    INFORMATION FOR SEQ ID NO:  13:
     SEQUENCE CHARACTERISTICS:
       LENGTH:  1721 base pairs
       TYPE:  nucleic acid
       STRANDEDNESS:  single
       TOPOLOGY:  linear
     MOLECULE TYPE:  cDNA
     FEATURE:
       NAME/KEY:  CDS
       LOCATION:  66..1274
US-08-206-188B-13


  Query Match             0.5%;  Score 16;  DB 3;  Length 1721;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2731 cacagccgcgggccag 2746
          ||||||||||||||||
Db    266 CACAGCCGCGGGCCAG 281


RESULT  95
PCT-US91-02714-13
  Sequence 13, Application PC/TUS9102714
   GENERAL INFORMATION:
     APPLICANT:  Wigler, Michael H.
;    APPLICANT:  Colicelli, John J.
;    TITLE OF INVENTION:  Cloning by Complementation and Related
;    TITLE OF INVENTION:  Processes
;    NUMBER OF SEQUENCES:  55
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  Two First National Plaza, 20 South Clark
;      STREET:  Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/02714
;      FILING DATE:  19910419
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/511,715
;      FILING DATE:  20-APR-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25447
;      REFERENCE/DOCKET NUMBER:  27805/30197
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 346-5750
;      TELEFAX:  (312) 984-9740
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1721 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  60..1274
PCT-US91-02714-13


  Query Match             0.5%;  Score 16;  DB 4;  Length 1721;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2731 cacagccgcgggccag 2746
          ||||||||||||||||
Db    266 CACAGCCGCGGGCCAG 281


RESULT  96
US-08-481-814A-2
; Sequence 2, Application US/08481814A
; Patent No. 5869040
;  GENERAL INFORMATION:
;    APPLICANT:  Hsu, Yen-Ming
;    TITLE OF INVENTION:  GENE THERAPY METHODS AND COMPOSITIONS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Biogen, Inc.
;      STREET:  14 Cambridge Center
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  United States of America
;      ZIP:  02142
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/481,814A
      FILING DATE:
      CLASSIFICATION:  424
    ATTORNEY/AGENT INFORMATION:
      NAME:  Kaplan, Warren A
      REFERENCE/DOCKET NUMBER:  A001
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  617-679-2000
      TELEFAX:  617-679-2838
  INFORMATION FOR SEQ ID NO:  2:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  1766 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  double
      TOPOLOGY:  linear
    MOLECULE TYPE:  cDNA to mRNA
    HYPOTHETICAL:  NO
    ANTI-SENSE:  NO
    ORIGINAL SOURCE:
      ORGANISM:  Homo sapiens
      CELL LINE:  HeLa
    FEATURE:
      NAME/KEY:  mat_peptide
      LOCATION:  429..1739
      OTHER INFORMATION:  /product= "E2F-2"
S-08-481-814A-2


  Query Match              0.5%;  Score 16;  DB 2;  Length 1766;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

y   2735 gccgcgggccaggagg 2750
       ||||||||||||||||
b    262 GCCGCGGGCCAGGAGG 277


RESULT  97
S-08-993-228-1
 Sequence 1, Application US/08993228
 Patent No. 5976838
  GENERAL INFORMATION:
    APPLICANT:  Jacobs, Kenneth
    APPLICANT:  McCoy, John M.
    APPLICANT:  LaVallie, Edward R.
    APPLICANT:  Racie, Lisa A.
    APPLICANT:  Merberg, David
    APPLICANT:  Treacy, Maurice
    APPLICANT:  Spaulding, Vikki
    APPLICANT:  Agostino, Michael J.
    TITLE OF INVENTION:  SECRETED PROTEINS AND POLYNUCLEOTIDES
    TITLE OF INVENTION:  ENCODING THEM
    NUMBER OF SEQUENCES:  33
    CORRESPONDENCE ADDRESS:
      ADDRESSEE:  Genetics Institute, Inc.
      STREET:  87 CambridgePark Drive
      CITY:  Cambridge
      STATE:  MA
      COUNTRY:  U.S.A.
      ZIP:  02140
    COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  US/08/993,228
      FILING DATE:
      CLASSIFICATION:
    ATTORNEY/AGENT INFORMATION:
      NAME:  Sprunger, Suzanne A.
      REGISTRATION NUMBER:  41,323
    TELECOMMUNICATION INFORMATION:
      TELEPHONE:  (617) 498-8284
      TELEFAX:  (617) 876-5851
  INFORMATION FOR SEQ ID NO:  1:
    SEQUENCE CHARACTERISTICS:
      LENGTH:  1790 base pairs
      TYPE:  nucleic acid
      STRANDEDNESS:  double
      TOPOLOGY:  linear
    MOLECULE TYPE:  cDNA
US-08-993-228-1


  Query Match              0.5%;  Score 16;  DB 2;  Length 1790;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1581 tctctgctactggact 1596
          ||||||||||||||||
Db   1235 TCTCTGCTACTGGACT 1250


RESULT  98
US-09-231-529-2
; Sequence 2, Application US/09231529
; Patent No. 6096308
;  GENERAL INFORMATION:
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  HUMAN PROTEIN KINASE AND KINASE INHIBITORS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/231,529
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/977,816
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0429 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1977 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  KIDNNOT25
;      CLONE:  3453694
US-09-231-529-2
```

```
  Query Match             0.5%;  Score 16;  DB 3;  Length 1977;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1462 cagccccagcagagaa 1477
          ||||||||||||||||
Db    419 CAGCCCCAGCAGAGAA 434


RESULT  99
US-09-255-911-1
; Sequence 1, Application US/09255911
; Patent No. 6013522
; GENERAL INFORMATION:
;  APPLICANT: Brett P. Monia
;  APPLICANT: Lex M. Cowsert
;  TITLE OF INVENTION: ANTISENSE MODULATION OF SMAD1 EXPRESSION
;  FILE REFERENCE: RTS-0040
;  CURRENT APPLICATION NUMBER: US/09/255,911
;  CURRENT FILING DATE: 1999-02-23
;  NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 1
;   LENGTH: 1990
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (433)..(1830)
US-09-255-911-1


  Query Match             0.5%;  Score 16;  DB 3;  Length 1990;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1077 atggccccagcatctg 1092
          ||||||||||||||||
Db   1831 atggccccagcatctg 1846


RESULT 100
US-08-619-554-7
; Sequence 7, Application US/08619554
; Patent No. 5821353
; GENERAL INFORMATION:
;    APPLICANT:  DOUGLAS, Cameron M.
;    APPLICANT:  CHREBET, Gary L.
;    APPLICANT:  CLEMAS, Joseph
;    APPLICANT:  EL-SHERBEINI, Mohammed
;    APPLICANT:  FOOR, Forrest
;    APPLICANT:  KAHN, Jennifer,
;    APPLICANT:  KELLY, Rosemarie, - PARENT, S.A.
;    APPLICANT:  MARRINAN, Jean, - RAMADAN, N.M.
;    APPLICANT:  MORIN, Nancy, - REGISTER, E.A
;    APPLICANT:  ONISHI, Janet, - SHEI, Gan-Ju
;    TITLE OF INVENTION:  DNA ENCODING 1,3 BETA-D GLUCAN
;    TITLE OF INVENTION:  SYNTHASE SUBUNITS
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  JOSEPH A. COPPOLA - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/619,554
;      FILING DATE:  01-AUG-1996
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COPPOLA, JOSEPH A
;      REGISTRATION NUMBER:  38,413
;      REFERENCE/DOCKET NUMBER:  19104PI
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  732-594-6734
;      TELEFAX:  732-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2069 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-619-554-7


  Query Match             0.5%;  Score 16;  DB 1;  Length 2069;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    919 agattttggctgaaga 934
          ||||||||||||||||
Db   1298 AGATTTTGGCTGAAGA 1313


Search completed: February 18, 2001, 14:23:44
Job time: 24889 sec
```

GenCore version 4.5

.M nucleic - nucleic search, using sw model

Run on: February 18, 2001, 05:22:53 ; Search time 2148.29 Seconds
(without alignments)
9648.673 Million cell updates/sec

Title: US-09-434-382-3
Perfect score: 2958
Sequence: 1 cgcgggcgtaggtgaccggc.....aataaagattgagtttgcaa 2958

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 7991742 seqs, 3503743858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: em_estba:*
42: em_estfun:*
43: em_esthum1:*
44: em_esthum2:*
45: em_esthum3:*
46: em_esthum4:*
47: em_esthum5:*
48: em_esthum6:*
49: em_esthum7:*
50: em_esthum8:*
51: em_esthum9:*
52: em_esthum10:*
53: em_esthum11:*
54: em_esthum12:*
55: em_esthum13:*
56: em_esthum14:*
57: em_esthum15:*
58: em_esthum16:*
59: em_esthum17:*
60: em_esthum18:*
61: em_esthum19:*
62: em_esthum20:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estov1:*
68: em_estov2:*
69: em_estpl1:*
70: em_estpl2:*
71: em_estpl3:*
72: em_estpl4:*
73: em_estpl5:*
74: em_estro1:*
75: em_estro2:*
76: em_estro3:*
77: em_estro4:*
78: em_estro5:*
79: em_estro6:*
80: em_estro7:*
81: em_estro8:*
82: em_estro9:*
83: em_estro10:*
84: em_estro11:*
85: em_estro12:*
86: em_estro13:*
87: gb_est41:*
88: gb_est42:*
89: gb_est43:*
90: gb_est44:*
91: gb_est45:*
92: gb_est46:*
93: gb_est47:*
94: gb_est48:*
95: gb_est49:*
96: gb_est50:*
97: gb_est51:*
98: gb_est52:*
99: gb_est53:*
100: gb_est54:*
101: gb_est55:*
102: gb_est56:*
103: gb_est57:*
104: gb_est67:*
105: gb_est68:*
106: gb_est69:*
107: gb_est70:*
108: gb_est71:*
109: gb_est72:*
110: gb_est73:*
111: gb_est74:*
112: em_esthum21:*
113: em_esthum22:*
114: em_esthum23:*
115: em_estom1:*
116: em_estom2:*

117: em_estpl6:*
118: em_estpl7:*
119: em_estpl8:*
120: em_estrol4:*
121: em_estrol5:*
122: em_estrol6:*
123: em_estrol7:*
124: em_estrol8:*
125: em_estrol9:*
126: gb_est58:*
127: gb_est59:*
128: gb_est60:*
129: gb_est61:*
130: gb_est62:*
131: gb_est63:*
132: gb_est64:*
133: gb_est65:*
134: gb_est66:*
135: gb_est75:*
136: gb_est76:*
137: gb_est77:*
138: gb_est78:*
139: gb_est79:*
140: gb_est80:*
141: gb_est81:*
142: gb_est82:*
143: gb_est83:*
144: gb_est84:*
145: gb_est85:*
146: gb_est86:*
147: gb_est87:*
148: gb_gss1:*
149: gb_gss2:*
150: gb_gss3:*
151: gb_gss4:*
152: em_gss1:*
153: em_gss2:*
154: em_gss3:*
155: em_gss4:*
156: gb_gss5:*
157: gb_gss6:*
158: em_gss5:*
159: em_gss6:*
160: gb_gss7:*
161: gb_gss8:*
162: gb_gss9:*
163: em_gss7:*
164: em_gss8:*
165: gb_gss11:*
166: gb_gss10:*
167: em_gss9:*
168: em_gss10:*
169: em_gss11:*
170: em_gss12:*
171: gb_gss12:*
172: gb_gss13:*
173: gb_gss14:*
174: gb_gss15:*
175: gb_gss16:*
176: gb_gss17:*
177: gb_gss18:*
178: gb_gss19:*
179: em_gss13:*
180: gb_gss20:*
181: gb_gss21:*
182: gb_gss22:*
183: gb_gss23:*
184: gb_gss24:*
185: em_gss14:*
186: em_gss15:*
187: em_gss16:*
188: em_gss17:*
189: em_gss18:*
190: gb_gss25:*
191: gb_gss26:*
192: gb_gss27:*
193: gb_gss28:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	676	22.9	676	106	BE260495	BE260495 601150702
c	2	614	20.8	975	105	BE250412	BE250412 600943455
	3	590	19.9	664	107	BE383336	BE383336 601298249
	4	590	19.9	692	107	BE382353	BE382353 601298656
	5	537	18.2	627	107	BE386924	BE386924 601274815
	6	528	17.8	761	135	BE795820	BE795820 601590856
	7	521	17.6	688	135	BE794311	BE794311 601591442
c	8	478	16.2	531	92	AW572950	AW572950 hf17h05.x
	9	465	15.7	823	110	BE619259	BE619259 601473130
	10	458	15.5	531	4	AA243700	AA243700 zr68g08.s
	11	451	15.2	451	17	AI200296	AI200296 qf86b12.x
	12	450	15.2	812	136	BE867512	BE867512 601443010
c	13	448	15.1	499	91	AW510825	AW510825 hd40b11.x
	14	447	15.1	938	106	BE260626	BE260626 601146116
	15	442	14.9	493	92	AW575677	AW575677 UI-HF-BM0
	16	438	14.8	612	106	BE304720	BE304720 601106236
	17	436	14.7	493	7	AA455121	AA455121 zx78c04.s
c	18	425	14.4	478	92	AW592601	AW592601 hf45a09.x
c	19	425	14.4	527	16	AI089646	AI089646 qb16g07.x
	20	421	14.2	421	20	AI468143	AI468143 tf92g05.x
	21	421	14.2	536	87	AW206103	AW206103 UI-H-BI1-
	22	404	13.7	404	88	AW304130	AW304130 xs13e05.x
	23	402	13.6	474	4	AA291670	AA291670 zt37d04.s
	24	400	13.5	992	135	BE747163	BE747163 601577254
	25	384	13.0	574	105	BE250309	BE250309 600943455
	26	380	12.8	431	5	AA310236	AA310236 EST181085
	27	373	12.6	949	135	BE744197	BE744197 601577168
	28	370	12.5	370	137	BE883616	BE883616 601508091
	29	368	12.4	745	137	BE900936	BE900936 601674206
	30	352	11.9	518	144	R87541	R87541 ym89b04.r1
c	31	351	11.9	855	146	W27286	W27286 28h1 Human
	32	347	11.7	790	137	BE892893	BE892893 601435738
c	33	343	11.6	396	12	AA811170	AA811170 ob42c03.s
	34	342	11.6	446	10	AA634909	AA634909 ab27h02.r
	35	342	11.6	452	10	AA679618	AA679618 ag72c12.s
	36	341	11.5	698	135	BE795434	BE795434 601592991
c	37	338	11.4	345	90	AW407520	AW407520 UI-HF-BM0
	38	334	11.3	677	135	BE742908	BE742908 601574609
	39	330	11.2	416	106	BE298273	BE298273 601118143
	40	329	11.1	940	135	BE743831	BE743831 601577742
c	41	325	11.0	531	27	AI937465	AI937465 wp77e01.x
	42	323	10.9	397	10	AA632118	AA632118 np66h03.s
	43	323	10.9	482	144	R55841	R55841 yg89d01.r1
c	44	315	10.6	366	4	AA233087	AA233087 zr68g08.r
	45	307	10.4	735	137	BE902696	BE902696 601677393
c	46	302	10.2	491	10	AA676661	AA676661 zj67h01.s
	47	301	10.2	480	144	R90875	R90875 yn01d02.r1
	48	301	10.2	511	95	AW771657	AW771657 hn59h06.x
	49	291	9.8	297	97	AW994476	AW994476 RC3-BN003
	50	291	9.8	404	17	AI174501	AI174501 an42e05.s
c	51	290	9.8	494	136	BE858252	BE858252 7g21a09.x
	52	287	9.7	414	105	BE243887	BE243887 TCBAP1E15
	53	286	9.7	355	6	AA353573	AA353573 EST61719
	54	285	9.6	397	146	W37486	W37486 zc10f03.s1
c	55	284	9.6	394	88	AW248468	AW248468 2820540.3
c	56	283	9.6	489	25	AI803400	AI803400 tc42f03.x
	57	282	9.5	307	145	T34216	T34216 EST64346 Hu
c	58	281	9.5	489	8	AA534478	AA534478 nf76f10.s
	59	276	9.3	278	95	AW806551	AW806551 IL0-ST000

	60	276	9.3	461	10	AA635046 AA635046 ab48b06.r
c	61	274	9.3	364	87	AW175581 AW175581 QV0-BT004
	62	273	9.2	457	14	AA994126 AA994126 ou38b06.s
	63	266	9.0	984	135	BE744876 BE744876 601576324
c	64	255	8.6	410	15	AI033108 AI033108 ow98g08.s
c	65	255	8.6	461	92	AW592223 AW592223 hf41a01.x
c	66	253	8.6	479	11	AA716607 AA716607 zg68g07.s
	67	250	8.5	501	5	AA311855 AA311855 EST182568
	68	248	8.4	299	147	Z17886 Z17886 HSDHEI030 S
c	69	247	8.4	517	27	AI991599 AI991599 ws18c04.x
	70	246	8.3	387	137	BE938229 BE938229 CM4-TN006
c	71	245	8.3	429	12	AA838624 AA838624 oe91f04.s
c	72	243	8.2	346	96	AW889463 AW889463 RC6-NT002
c	73	238	8.0	414	141	H22087 H22087 yl34c04.r1
	74	236	8.0	691	107	BE409312 BE409312 601300940
c	75	235	7.9	439	17	AI201492 AI201492 qs74b03.x
c	76	233	7.9	233	14	AF188525 AF188525
c	77	233	7.9	394	141	H14462 H14462 yl25f04.r1
c	78	232	7.8	865	8	AA522537 AA522537 ni38e08.s
c	79	228	7.7	282	13	AA928608 AA928608 om75b03.s
	80	228	7.7	282	145	T34024 T34024 EST61387 Hu
c	81	227	7.7	433	19	AI357786 AI357786 qu98d07.x
	82	224	7.6	872	110	BE615669 BE615669 601279347
	83	221	7.5	291	4	AA235532 AA235532 zt29g11.s
c	84	220	7.4	477	25	AI804749 AI804749 tu42d02.x
	85	218	7.4	228	88	AW296524 UI-H-BW0-
	86	218	7.4	448	142	N36229 N36229 yy30c04.s1
c	87	214	7.2	316	91	AW511765 AW511765 xu76f03.x
	88	211	7.1	249	88	AW247380 AW247380 2820640.5
c	89	209	7.1	650	89	AW378247 AW378247 RC1-HT021
c	90	207	7.0	219	8	AA504146 AA504146 aa59e06.s
c	91	206	7.0	422	146	W37591 W37591 zc10f03.r1
	92	199	6.7	472	144	R51138 R51138 yg71c08.r1
c	93	198	6.7	416	16	AI141263 AI141263 qa46h05.s
	94	193	6.5	394	6	AA346268 AA346268 EST52407
	95	184	6.2	276	6	AA378232 AA378232 EST90926
	96	182	6.2	290	147	Z44544 Z44544 HSC23E061 n
c	97	182	6.2	424	141	H03318 H03318 yj47e10.s1
c	98	180	6.1	482	15	AI033342 AI033342 ox02d12.s
c	99	179	6.1	577	110	BE619874 BE619874 601473130
	100	178	6.0	376	145	T72963 T72963 yc65b06.r1

ALIGNMENTS

```
RESULT   1
BE260495
LOCUS       BE260495      676 bp    mRNA            EST       13-JUL-2000
DEFINITION  601150702F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3503184 5',
            mRNA sequence.
ACCESSION   BE260495
VERSION     BE260495.1  GI:9131807
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 676)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM176  row: d  column: 01
            High quality sequence stop: 672.

FEATURES             Location/Qualifiers
     source          1. .676
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3503184"
                     /clone_lib="NIH_MGC_19"
                     /tissue_type="neuroblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and  Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
BASE COUNT      154 a    207 c    176 g    139 t
ORIGIN


  Query Match             22.9%;  Score 676;  DB 106;  Length 676;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  676;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1143 acccagcacttggtcctgaatgagaactgtgcctcagttcacaaccttcgcagccacaag 1202
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTTCGCAGCCACAAG 60

Qy    1203 attcaaacccagctcaacctcatccacccggacatcttccccctgctcaccagtttccgc 1262
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTCACCAGTTTCCGC 120

Qy    1263 tgtaagaaggagggccccaccctcagtgtgcccatggttcagggtgaatgcctcctcaag 1322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAATGCCTCCTCAAG 180

Qy    1323 taccagctccgtcccaggagggagtggcagagggatgccattattacttgcaatcctgag 1382
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACTTGCAATCCTGAG 240

Qy    1383 gaattcatagttgaggcgctgcagcttcccaacttccagcagagcgtgcaggagtacagg 1442
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTGCAGGAGTACAGG 300

Qy    1443 aggagtgcgcaggacggcccagccccagcagagaaaagaagtcagtacccagaaatcatc 1502
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTACCCAGAAATCATC 360

Qy    1503 ttccttggaacagggtctgccatcccgatgaagattcgaaatgtcagtgccacacttgtc 1562
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGTGCCACACTTGTC 420

Qy    1563 aacataagccccgacacgtctctgctactggactgtggtgagggcacatttgggcagctg 1622
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 AACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACATTTGGGCAGCTG 480

Qy    1623 tgccgtcattacggagaccaggtggacagggtcctgggcaccctggctgctgtgtttgtg 1682
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCTGCTGTGTTTGTG 540

Qy    1683 tcccacctgcacgcagatcaccacacgggcttgccaagtatcttgctgcagagagaacgc 1742
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 TCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTGCAGAGAGAACGC 600

Qy    1743 gccttggcatctttgggaaagccgcttcacccttgctggtggttgccccccaaccagctc 1802
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCCCCCAACCAGCTC 660

Qy    1803 aaagcctggctccagc 1818
           ||||||||||||||||
Db     661 AAAGCCTGGCTCCAGC 676
```

```
RESULT   2
BE250412/c
LOCUS       BE250412      975 bp    mRNA            EST       13-JUL-2000
DEFINITION  60094345ST1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2960077 3',
            mRNA sequence.
ACCESSION   BE250412
VERSION     BE250412.1  GI:9120523
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 975)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM52  row: n  column: 14
            High quality sequence start: 22
            High quality sequence stop: 756.
FEATURES             Location/Qualifiers
     source          1. .975
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2960077"
                     /clone_lib="NIH_MGC_17"
                     /tissue_type="rhabdomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
                     Site_2: XhoI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Size-selected >500bp
                     for average insert size 1.8kb.  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      185 a    296 c    258 g    236 t
ORIGIN


  Query Match           20.8%;  Score 614;  DB 105;  Length 975;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  614;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2197 acgcggagttcattatgctgaaccacttcagccagcgctatgccaaggtccccctcttca 2256
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     711 ACGCGGAGTTCATTATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCA 652

Qy    2257 gccccaacttcagcgagaaagtgggagttgcctttgaccacatgaaggtctgctttggag 2316
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     651 GCCCCAACTTCAGCGAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAG 592

Qy    2317 actttccaacaatgcccaagctgattcccccactgaaagccctgtttgctggcgacatcg 2376
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     591 ACTTTCCAACAATGCCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCG 532

Qy    2377 aggagatggaggagcgcagggagaagcgggagctgcggcaggtgcgggcggccctcctgt 2436
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     531 AGGAGATGGAGGAGCGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGT 472

Qy    2437 ccagggagctggcaggcggcctggaggatggggagcctcagcagaagcgggcccacacag 2496
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     471 CCAGGGAGCTGGCAGGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAG 412

Qy    2497 aggagccacaggccaagaaggtcagagcccagtgaagatctgggagaccctgaactcaga 2556
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     411 AGGAGCCACAGGCCAAGAAGGTCAGAGCCCAGTGAAGATCTGGGAGACCCTGAACTCAGA 352

Qy    2557 aggctgtgtgtcttctgccccacgcacgcacccgtatctgccctccttgctggtagaagc 2616
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     351 AGGCTGTGTGTCTTCTGCCCCACGCACGCACCCGTATCTGCCCTCCTTGCTGGTAGAAGC 292

Qy    2617 tgaagagcacggtccccccaggaggcagctcaggataggtggtatggagctgtgccgaggc 2676
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     291 TGAAGAGCACGGTCCCCCAGGAGGCAGCTCAGGATAGGTGGTATGGAGCTGTGCCGAGGC 232

Qy    2677 ttgggctcccacataagcactagtctatagatgcctcttaggactggtgcctggcacagc 2736
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     231 TTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTTAGGACTGGTGCCTGGCACAGC 172

Qy    2737 cgcgggccaggaggctgccacacggaagcaagcagatgaactaatttcatttcaaggcag 2796
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     171 CGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGAACTAATTTCATTTCAAGGCAG 112

Qy    2797 tttttaaagaagtc 2810
           ||||||||||||||
Db     111 TTTTTAAAGAAGTC 98


RESULT   3
BE383336
LOCUS       BE383336      664 bp    mRNA            EST       21-JUL-2000
DEFINITION  601298249F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3628308 5',
            mRNA sequence.
ACCESSION   BE383336
VERSION     BE383336.1  GI:9328701
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 664)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM313  row: a  column: 13
            High quality sequence stop: 662.
FEATURES             Location/Qualifiers
     source          1. .664
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3628308"
                     /clone_lib="NIH_MGC_19"
                     /tissue_type="neuroblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and  Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
BASE COUNT      154 a    199 c    174 g    137 t
ORIGIN


  Query Match           19.9%;  Score 590;  DB 107;  Length 664;
  Best Local Similarity  100.0%;  Pred. No. 0;
```

```
  Matches  590;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1143 acccagcacttggtcctgaatgagaactgtgcctcagttcacaaccttcgcagccacaag 1202
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTTCGCAGCCACAAG 60

Qy   1203 attcaaacccagctcaacctcatccacccggacatcttccccctgctcaccagtttccgc 1262
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTCACCAGTTTCCGC 120

Qy   1263 tgtaagaaggagggccccaccctcagtgtgcccatggttcagggtgaatgcctcctcaag 1322
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAATGCCTCCTCAAG 180

Qy   1323 taccagctccgtcccaggagggagtggcagagggatgccattattacttgcaatcctgag 1382
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACTTGCAATCCTGAG 240

Qy   1383 gaattcatagttgaggcgctgcagcttcccaacttccagcagagcgtgcaggagtacagg 1442
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTGCAGGAGTACAGG 300

Qy   1443 aggagtgcgcaggacggcccagccccagcagagaaaagaagtcagtacccagaaatcatc 1502
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTACCCAGAAATCATC 360

Qy   1503 ttccttggaacagggtctgccatcccgatgaagattcgaaatgtcagtgccacacttgtc 1562
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGTGCCACACTTGTC 420

Qy   1563 aacataagccccgacacgtctctgctactggactgtggtgagggcacatttgggcagctg 1622
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACATTTGGGCAGCTG 480

Qy   1623 tgccgtcattacggagaccaggtggacagggtcctgggcaccctggctgctgtgtttgtg 1682
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCTGCTGTGTTTGTG 540

Qy   1683 tcccacctgcacgcagatcaccacacgggcttgccaagtatcttgctgca 1732
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTGCA 590


RESULT   4
BE382353
LOCUS       BE382353      692 bp    mRNA            EST       21-JUL-2000
DEFINITION  601298656F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3629028 5',
            mRNA sequence.
ACCESSION   BE382353
VERSION     BE382353.1  GI:9327718
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 692)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM314  row: o  column: 13
            High quality sequence stop: 600.
FEATURES             Location/Qualifiers
     source          1. .692
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3629028"
                     /clone_lib="NIH_MGC_19"
                     /tissue_type="neuroblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and  Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
BASE COUNT      152 a    193 c    197 g    150 t
ORIGIN


  Query Match             19.9%;  Score 590;  DB 107;  Length 692;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  590;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1525 tcccgatgaagattcgaaatgtcagtgccacacttgtcaacataagccccgacacgtctc 1584
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TCCCGATGAAGATTCGAAATGTCAGTGCCACACTTGTCAACATAAGCCCCGACACGTCTC 60

Qy   1585 tgctactggactgtggtgagggcacatttgggcagctgtgccgtcattacggagaccagg 1644
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGCTACTGGACTGTGGTGAGGGCACATTTGGGCAGCTGTGCCGTCATTACGGAGACCAGG 120

Qy   1645 tggacagggtcctgggcaccctggctgctgtgtttgtgtcccacctgcacgcagatcacc 1704
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGGACAGGGTCCTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACC 180

Qy   1705 acacgggcttgccaagtatcttgctgcagagagaacgcgccttggcatctttgggaaagc 1764
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACACGGGCTTGCCAAGTATCTTGCTGCAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGC 240

Qy   1765 cgcttcacccttgctggtggttgcccccaaccagctcaaagcctggctccagcagtacc 1824
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CGCTTCACCCTTTGCTGGTGGTTGCCCCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACC 300

Qy   1825 acaaccagtgccaggaggtcctgcaccacatcagtatgattcctgccaaatgccttcagg 1884
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ACAACCAGTGCCAGGAGGTCCTGCACCACATCAGTATGATTCCTGCCAAATGCCTTCAGG 360

Qy   1885 aaggggctgagatctccagtcctgcagtggaaagattgatcagttcgctgttgcgaacat 1944
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAGGGGCTGAGATCTCCAGTCCTGCAGTGGAAAGATTGATCAGTTCGCTGTTGCGAACAT 420

Qy   1945 gtgatttggaagagtttcagacctgtctggtgcggcactgcaagcatgcgtttggctgtg 2004
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTGATTTGGAAGAGTTTCAGACCTGTCTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTG 480

Qy   2005 cgctggtgcacacctctggctggaaagtggtctattccggggacaccatgccctgcgagg 2064
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CGCTGGTGCACACCTCTGGCTGGAAAGTGGTCTATTCCGGGGACACCATGCCCTGCGAGG 540

Qy   2065 ctctggtccggatggggaaagatgccaccctcctgatacatgaagccacc 2114
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CTCTGGTCCGGATGGGGAAAGATGCCACCCTCCTGATACATGAAGCCACC 590


RESULT   5
BE386924
LOCUS       BE386924      627 bp    mRNA            EST       21-JUL-2000
DEFINITION  601274815F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3615939 5',
            mRNA sequence.
ACCESSION   BE386924
VERSION     BE386924.1  GI:9332387
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
EFERENCE    1  (bases 1 to 627)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
OMMENT      Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM280  row: n  column: 04
            High quality sequence start: 18
            High quality sequence stop: 611.
EATURES             Location/Qualifiers
     source         1. .627
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:3615939"
                    /clone_lib="NIH_MGC_20"
                    /tissue_type="melanotic melanoma"
                    /lab_host="DH10B (phage-resistant)"
                    /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                    EcoRI; cDNA made by oligo-dT priming.  Directionally
                    cloned into EcoRI/XhoI sites using the following 5'
                    adaptor: GGCACGAG(G).  Size-selected >500bp for average
                    insert size 1.8kb.  Library constructed by Ling Hong in
                    the laboratory of Gerald M. Rubin (University of
                    California, Berkeley) using  ZAP-cDNA synthesis kit
                    (Stratagene) and Superscript II RT (Life Technologies)."
ASE COUNT      147 a    175 c    199 g    106 t
RIGIN
```

```
  Query Match           18.2%;  Score 537;  DB 107;  Length 627;
  Best Local Similarity  100.0%;  Pred. No. 7.1e-273;
  Matches  537;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

y   2145 gaaaagacacacagcacaacgtcccaagccatcagcgtggggatgcggatgaacgcggag 2204
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     16 GAAAAGACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAG 75

y   2205 ttcattatgctgaaccacttcagccagcgctatgccaaggtccccctcttcagccccaac 2264
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     76 TTCATTATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAAC 135

y   2265 ttcagcgagaaagtgggagttgcctttgaccacatgaaggtctgctttggagactttcca 2324
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    136 TTCAGCGAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCA 195

y   2325 acaatgcccaagctgattcccccactgaaagccctgtttgctggcgacatcgaggagatg 2384
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    196 ACAATGCCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATG 255

y   2385 gaggagcgcagggagaagcgggagctgcggcaggtgcgggcggccctcctgtccagggag 2444
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    256 GAGGAGCGCAGGGAGAAGCGGGAGCTGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGGGAG 315

y   2445 ctggcaggcggcctggaggatggggagcctcagcagaagcgggcccacacagaggagcca 2504
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    316 CTGGCAGGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCA 375

y   2505 caggccaagaaggtcagagcccagtgaagatctgggagaccctgaactcagaaggctgtg 2564
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    376 CAGGCCAAGAAGGTCAGAGCCCAGTGAAGATCTGGGAGACCCTGAACTCAGAAGGCTGTG 435

y   2565 tgtcttctgccccacgcacgcacccgtatctgccctccttgctggtagaagctgaagagc 2624
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    436 TGTCTTCTGCCCCACGCACGCACCCGTATCTGCCCTCCTTGCTGGTAGAAGCTGAAGAGC 495
```

```
Qy   2625 acggtcccccaggaggcagctcaggataggtggtatggagctgtgccgaggcttggg 2681
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    496 ACGGTCCCCCAGGAGGCAGCTCAGGATAGGTGGTATGGAGCTGTGCCGAGGCTTGGG 552


RESULT   6
BE795820
LOCUS       BE795820      761 bp    mRNA            EST       20-SEP-2000
DEFINITION  601590856F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3945085 5',
            mRNA sequence.
ACCESSION   BE795820
VERSION     BE795820.1  GI:10217018
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 761)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM802  row: p  column: 14
            High quality sequence stop: 761.
FEATURES            Location/Qualifiers
     source         1. .761
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:3945085"
                    /clone_lib="NIH_MGC_7"
                    /tissue_type="small cell carcinoma"
                    /cell_line="MGC3"
                    /lab_host="DH10B (phage-resistant)"
                    /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                    EcoRI; cDNA made by oligo-dT priming.  Directionally
                    cloned into EcoRI/XhoI sites using the following 5'
                    adaptor: GGCACGAG(G).  Size-selected >500bp for average
                    insert size 1.8kb.  Library constructed by Ling Hong in
                    the laboratory of Gerald M. Rubin (University of
                    California, Berkeley) using  ZAP-cDNA synthesis kit
                    (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      172 a    206 c    219 g    164 t
ORIGIN
```

```
  Query Match           17.8%;  Score 528;  DB 135;  Length 761;
  Best Local Similarity  99.7%;  Pred. No. 4.2e-268;
  Matches  698;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy   1529 gatgaagattcgaaatgtcagtgccacacttgtcaacataagccccgacacgtctctgct 1588
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GATGAAGATTCGAAATGTCAGTGCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCT 60

Qy   1589 actggactgtggtgagggcacatttgggcagctgtgccgtcattacggagaccaggtgga 1648
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db     61 ACTGGACTGTGGTGAGGGCACGTTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGA 120

Qy   1649 cagggtcctgggcaccctggctgctgtgtttgtgtcccacctgcacgcagatcaccacac 1708
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAGGGTCCTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACAC 180

Qy   1709 gggcttgccaagtatcttgctgcagagagaacgcgccttggcatctttgggaaagccgct 1768
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGGCTTGCCAAGTATCTTGCTGCAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCT 240
```

```
Qy   1769 tcacccttttgctggtggttgcccccaaccagctcaaagcctggctccagcagtaccacaa 1828
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCACCCTTTGCTGGTGGTTGCCCCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAA 300

Qy   1829 ccagtgccaggaggtcctgcaccacatcagtatgattcctgccaaatgccttcaggaagg 1888
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CCAGTGCCAGGAGGTCCTGCACCACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGG 360

Qy   1889 ggctgagatctccagtcctgcagtggaaagattgatcagttcgctgttgcgaacatgtga 1948
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGCTGAGATCTCCAGTCCTGCAGTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGA 420

Qy   1949 tttggaagagtttcagacctgtctggtgcggcactgcaagcatgcgtttggctgtgcgct 2008
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 -TTGGAAGAGTTTCAGACCTGTCTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCT 479

Qy   2009 ggtgcacacctctggctggaaagtggtctattccggggacaccatgccctgcgaggctct 2068
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 GGTGCACACCTCTGGCTGGAAAGTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCT 539

Qy   2069 ggtccggatggggaaagatgccaccctcctgatacatgaagccaccctggaagatggttt 2128
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 GGTCCGGATGGGGAAAGATGCCACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTT 599

Qy   2129 ggaagaggaagcagtggaaaagacacacagcacaacgtcccaagccatcagcgtggggat 2188
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    600 GGAAGAGGAAGCAGTGGAAAAGACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGAT 659

Qy   2189 gcggatgaacgcggagttcattatgctgaaccacttcagc 2228
          ||||||||||||||||||||||||||||||||||||||||
Db    660 GCGGATGAACGCGGAGTTCATTATGCTGAACCACTTCAGC 699
```

RESULT 7
BE794311
LOCUS BE794311 688 bp mRNA EST 20-SEP-2000
DEFINITION 601591442F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3945538 5', mRNA sequence.
ACCESSION BE794311
VERSION BE794311.1 GI:10215496
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 688)
AUTHORS NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM804 row: c column: 11
High quality sequence stop: 688.
FEATURES Location/Qualifiers
source 1. .688
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3945538"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 169 a 190 c 185 g 144 t
ORIGIN

Query Match 17.6%; Score 521; DB 135; Length 688;
Best Local Similarity 99.8%; Pred. No. 2.1e-264;
Matches 641; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy    820 tcaaagcaaaggagatgggcctcccagttgggacagctgccatcgctcccatcattgctg 879
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TCAAAGCAAAGGAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTG 60

Qy    880 ctgtcaaggacgggaaaagcatcactcatgaaggaagagagattttggctgaagagctgt 939
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGTCAAGGACGGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGT 120

Qy    940 gtactcctccagatcctggtgctgcttttgtggtggtagaatgtccagatgaaagcttca 999
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTACTCCTCCAGATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCA 180

Qy   1000 ttcaacccatctgtgagaatgccacctttcagaggtaccaaggaaaggcagatgcccccg 1059
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTCAACCCATCTGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCG 240

Qy   1060 tggccttggtggttcacatggccccagcatctgtgcttgtggacagcaggtaccagcagt 1119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TGGCCTTGGTGGTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGT 300

Qy   1120 ggatggagaggtttgggcctgacacccagcacttggtcctgaatgagaactgtgcctcag 1179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGATGGAGAGGTTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAG 360

Qy   1180 ttcacaaccttcgcagccacaagattcaaacccagctcaacctcatccacccggacatct 1239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTCACAACCTTCGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCT 420

Qy   1240 tccccctgctcaccagtttccgctgtaagaaggagggccccacctcagtgtgcccatgg 1299
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db    421 TCCCCCTGCTCACCAGTTTCCGCTGTAAGAAGGAGGGCCCCA-CCTCAGTGTGCCCATGG 479

Qy   1300 ttcagggtgaatgcctcctcaagtaccagctccgtcccaggagggagtggcagagggatg 1359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 TTCAGGGTGAATGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATG 539

Qy   1360 ccattattacttgcaatcctgaggaattcatagttgaggcgctgcagcttcccaacttcc 1419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 CCATTATTACTTGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCC 599

Qy   1420 agcagagcgtgcaggagtacaggaggagtgcgcaggacggcc 1461
          ||||||||||||||||||||||||||||||||||||||||
Db    600 AGCAGAGCGTGCAGGAGTACAGGAGGAGTGCGCAGGACGGCC 641
```

RESULT 8
AW572950/c
LOCUS AW572950 531 bp mRNA EST 13-MAR-2000
DEFINITION hf17h05.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2932185 3', mRNA sequence.
ACCESSION AW572950
VERSION AW572950.1 GI:7237683
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 531)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

```
  JOURNAL   Unpublished (1997)
OMMENT      Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 459.
EATURES             Location/Qualifiers
     source         1. .531
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:2932185"
                    /clone_lib="Soares_NFL_T_GBC_S1"
                    /lab_host="DH10B"
                    /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                    a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                    Equal amounts of plasmid DNA from three normalized
                    libraries (fetal lung NbHL19W, testis NHT, and B-cell
                    NCI_CGAP_GCB1) were mixed, and ss circles were made in
                    vitro. Following HAP purification, this DNA was used as
                    tracer in a subtractive hybridization reaction. The driver
                    was PCR-amplified cDNAs from pools of 5,000 clones made
                    from the same 3 libraries. The pools consisted of
                    I.M.A.G.E. clones 297480-302087, 682632-687239,
                    726408-728711, and 729096-731399. Subtraction by Bento
                    Soares and M. Fatima Bonaldo. "
ASE COUNT       99 a    154 c    136 g    142 t
RIGIN

  Query Match             16.2%;  Score 478;  DB 92;  Length 531;
  Best Local Similarity   99.8%;  Pred. No. 1.3e-241;
  Matches  528;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   2430 ctcctgtccagggagctggcaggcggcctggaggatggggagcctcagcagaagcgggcc 2489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    531 CTCCTGTCCAGGGAGCTGGCAGGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCC 472

Qy   2490 cacacagaggagccacaggccaagaaggtcagagcccagtgaagatctgggagaccctga 2549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    471 CACACAGAGGAGCCACAGGCCAAGAAGGTCAGAGCCCAGTGAAGATCTGGGAGACCCTGA 412

Qy   2550 actcagaaggctgtgtgtcttctgccccacgcacgcacccgtatctgccctccttgctgg 2609
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    411 ACTCAGAAGGCTGTGTGTCTTCTGCCCCACGCACGCACCCGTATCTGCCCTCCTTGCTGG 352

Qy   2610 tagaagctgaagagcacggtcccccaggaggcagctcaggataggtggtatggagctgtg 2669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    351 TAGAAGCTGAAGAGCACGGTCCCCCAGGAGGCAGCTCAGGATAGGTGGTATGGAGCTGTG 292

Qy   2670 ccgaggcttgggctcccacataagcactagtctatagatgcctcttaggactggtgcctg 2729
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    291 CCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTTAGGACTGGTGCCTG 232

Qy   2730 gcacagccgcgggccaggaggctgccacacggaagcaagcagatgaactaatttcatttc 2789
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    231 GCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGAACTAATTTCATTTC 172

Qy   2790 aaggcagtttttaaagaagtcttggaaacagacggcggcacctttcctctaatccagcaa 2849
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    171 AAGGCAGTTTTTAAAGAAGTCATGGAAACAGACGGCGGCACCTTTCCTCTAATCCAGCAA 112

Qy   2850 agtgattccctgcacaccagagacaagcagagtaacaggatcagtgggtctaagtgtccg 2909
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    111 AGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTGGGTCTAAGTGTCCG 52

Qy   2910 agacttaacgaaaatagtatttcagctgcaataaagattgagtttgcaa 2958
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db     51 AGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTGCAA 3

ESULT   9
```

```
BE619259
LOCUS       BE619259      823 bp    mRNA            EST       24-AUG-2000
DEFINITION  601473130F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3876223 5',
            mRNA sequence.
ACCESSION   BE619259
VERSION     BE619259.1  GI:9890197
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 823)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM623  row: k  column: 08
            High quality sequence stop: 695.
FEATURES            Location/Qualifiers
     source         1. .823
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:3876223"
                    /clone_lib="NIH_MGC_68"
                    /tissue_type="large cell carcinoma"
                    /lab_host="DH10B (phage-resistant)"
                    /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                    Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                    Average insert size 1.8 kb.   Library constructed by Life
                    Technologies. "
BASE COUNT      185 a    219 c    259 g    160 t
ORIGIN

  Query Match             15.7%;  Score 465;  DB 110;  Length 823;
  Best Local Similarity   100.0%;  Pred. No. 1e-234;
  Matches  465;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1678 ttgtgtcccacctgcacgcagatcaccacacgggcttgccaagtatcttgctgcagagag 1737
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     75 TTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTGCAGAGAG 134

Qy   1738 aacgcgccttggcatctttgggaaagccgcttcacccttgctggtggttgccccccaacc 1797
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 AACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCCCCCAACC 194

Qy   1798 agctcaaagcctggctccagcagtaccacaaccagtgccaggaggtcctgcaccacatca 1857
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 AGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCACCACATCA 254

Qy   1858 gtatgattcctgccaaatgccttcaggaaggggctgagatctccagtcctgcagtggaaa 1917
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255 GTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCAGTGGAAA 314

Qy   1918 gattgatcagttcgctgttgcgaacatgtgatttggaagagtttcagacctgtctggtgc 1977
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 GATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGTCTGGTGC 374

Qy   1978 ggcactgcaagcatgcgtttggctgtgcgctggtgcacacctctggctggaaagtggtct 2037
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 GGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAAGTGGTCT 434

Qy   2038 attccggggacaccatgccctgcgaggctctggtccggatggggaaagatgccaccctcc 2097
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    435 ATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCCACCCTCC 494

Qy   2098 tgatacatgaagccaccctggaagatggtttggaagaggaagcag 2142
          |||||||||||||||||||||||||||||||||||||||||||||
Db    495 TGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAG 539


RESULT 10
AA243700
LOCUS       AA243700      531 bp    mRNA            EST       07-MAR-1997
DEFINITION  zr68g08.s1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:668606 3'
            similar to SW:YK59_YEAST P36159 HYPOTHETICAL 96.8 KD PROTEIN IN
            SIS2-MTD1 INTERGENIC REGION. ;, mRNA sequence.
ACCESSION   AA243700
VERSION     AA243700.1  GI:1874492
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 531)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: similarity on wrong strand
            Possible reversed clone: polyT not found
            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 466.
FEATURES             Location/Qualifiers
     source          1. .531
                     /organism="Homo sapiens"
                     /db_xref="GDB:5562573"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:668606"
                     /clone_lib="Soares_NhHMPu_S1"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
BASE COUNT      144 a    137 c    143 g    107 t
ORIGIN


  Query Match             15.5%;  Score 458;  DB 4;  Length 531;
  Best Local Similarity   100.0%;  Pred. No. 5e-231;
  Matches  458;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    251 ccgggactcgggcgccgcgctctacgtcttctccgagttcaaccggtatctcttcaactg 310
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     74 CCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTCAACCGGTATCTCTTCAACTG 133

Qy    311 tggagaaggcgttcagagactcatgcaggagcacaagttaaaggttgctcgcctggacaa 370
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    134 TGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTAAAGGTTGCTCGCCTGGACAA 193

Qy    371 catattcctgacacgaatgcactggtctaatgttgggggcttaagtggaatgattcttac 430
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    194 CATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGCTTAAGTGGAATGATTCTTAC 253

Qy    431 tttaaaggaaaccgggcttccaaagtgtgtactttctggacctccacaactggaaaaata 490
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    254 TTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATA 313

Qy    491 cctcgaagcaatcaaaatattttctggtccattgaaaggaatagaactggctgtgcggcc 550
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    314 CCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGAATAGAACTGGCTGTGCGGCC 373

Qy    551 ccactctgccccagaatacgaggatgaaaccatgacagtttaccagatccccatacacag 610
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    374 CCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTTTACCAGATCCCCATACACAG 433

Qy    611 tgaacagaggaggggaaagcaccaaccatggcagagtccagaaaggcctctcagcaggct 670
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    434 TGAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGTCCAGAAAGGCCTCTCAGCAGGCT 493

Qy    671 cagtccagagcgatcttcagactccgagtcgaatgaaa 708
          |||||||||||||||||||||||||||||||||||||
Db    494 CAGTCCAGAGCGATCTTCAGACTCCGAGTCGAATGAAA 531


RESULT 11
AI200296
LOCUS       AI200296      451 bp    mRNA            EST       14-OCT-1998
DEFINITION  qf86b12.x1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone
            IMAGE:1756895 3' similar to SW:YK59_YEAST P36159 HYPOTHETICAL 96.8
            KD PROTEIN IN SIS2-MTD1 INTERGENIC REGION. ;, mRNA sequence.
ACCESSION   AI200296
VERSION     AI200296.1  GI:3752902
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 451)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 442.
FEATURES             Location/Qualifiers
     source          1. .451
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1756895"
                     /clone_lib="Soares_fetal_lung_NbHL19W"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCAATTTTTTTTTTTTTTTTT-3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. This library was constructed
                     from the same fetus as the fetal heart library, Soares
                     fetal heart NbHH19W."
```

```
ASE COUNT      110 a    120 c    128 g     93 t
RIGIN


  Query Match             15.2%;  Score 451;  DB 17;  Length 451;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-227;
  Matches  451;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

y    140 gcggccgcgcaaggacccgctgcggcacctgcgcacgcgagagaagcgcggaccgtcggg 199
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      1 GCGGCCGCGCAAGGACCCGCTGCGGCACCTGCGCACGCGAGAGAAGCGCGGACCGTCGGG 60

y    200 gtgctccggcggcccaaacaccgtgtacctgcaggtggtggcagcgggtagccgggactc 259
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     61 GTGCTCCGGCGGCCCAAACACCGTGTACCTGCAGGTGGTGGCAGCGGGTAGCCGGGACTC 120

y    260 gggcgccgcgctctacgtcttctccgagttcaaccggtatctcttcaactgtggagaagg 319
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    121 GGGCGCCGCGCTCTACGTCTTCTCCGAGTTCAACCGGTATCTCTTCAACTGTGGAGAAGG 180

y    320 cgttcagagactcatgcaggagcacaagttaaaggttgctcgcctggacaacatattcct 379
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    181 CGTTCAGAGACTCATGCAGGAGCACAAGTTAAAGGTTGCTCGCCTGGACAACATATTCCT 240

y    380 gacacgaatgcactggtctaatgttgggggcttaagtggaatgattcttacttttaaagga 439
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    241 GACACGAATGCACTGGTCTAATGTTGGGGGCTTAAGTGGAATGATTCTTACTTTAAAGGA 300

y    440 aaccgggcttccaaagtgtgtactttctggacctccacaactggaaaaatacctcgaagc 499
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    301 AACCGGGCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATACCTCGAAGC 360

y    500 aatcaaaatattttctggtccattgaaaggaatagaactggctgtgcggccccactctgc 559
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    361 AATCAAAATATTTTCTGGTCCATTGAAAGGAATAGAACTGGCTGTGCGGCCCCACTCTGC 420

y    560 cccagaatacgaggatgaaaccatgacagtt 590
        |||||||||||||||||||||||||||||||
b    421 CCCAGAATACGAGGATGAAACCATGACAGTT 451


ESULT  12
E867512
OCUS       BE867512      812 bp    mRNA            EST       27-SEP-2000
EFINITION  601443010F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3847226 5',
           mRNA sequence.
CCESSION   BE867512
ERSION     BE867512.1  GI:10316288
EYWORDS    EST.
OURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
EFERENCE   1  (bases 1 to 812)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
OMMENT     Contact: Robert Strausberg, Ph.D.
           Tel: (301) 496-1550
           Email: Robert_Strausberg@nih.gov
           Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLCM548  row: c  column: 03
           High quality sequence stop: 686.
EATURES             Location/Qualifiers
     source         1. .812
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:3847226"
                    /clone_lib="NIH_MGC_65"
                    /tissue_type="adenocarcinoma"
                    /lab_host="DH10B (phage-resistant)"
                    /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
                    Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                    Average insert size 1.8 kb.  Library constructed by Life
                    Technologies. "
BASE COUNT      200 a    216 c    220 g    176 t
ORIGIN


  Query Match             15.2%;  Score 450;  DB 136;  Length 812;
  Best Local Similarity   99.8%;  Pred. No. 9e-227;
  Matches  570;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy    774 gctttcatctgtaagcttcacttaaagagaggaaacttcttggtgctcaaagcaaaggag 833
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAGCAAAGGAG 60

Qy    834 atgggcctcccagttgggacagctgccatcgctcccatcattgctgctgtcaaggacggg 893
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAGGACGGG 120

Qy    894 aaaagcatcactcatgaaggaagagagattttggctgaagagctgtgtactcctccagat 953
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCAGAT 180

Qy    954 cctggtgctgcttttgtggtggtagaatgtccagatgaaagcttcattcaacccatctgt 1013
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCTGGTGCTGC-TTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATCTGT 239

Qy   1014 gagaatgccacctttcagaggtaccaaggaaaggcagatgcccccgtggccttggtggtt 1073
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 GAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTGGTT 299

Qy   1074 cacatggccccagcatctgtgcttgtggacagcaggtaccagcagtggatggagaggttt 1133
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 CACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGGTTT 359

Qy   1134 gggcctgacacccagcacttggtcctgaatgagaactgtgcctcagttcacaaccttcgc 1193
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 GGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTTCGC 419

Qy   1194 agccacaagattcaaacccagctcaacctcatccacccggacatcttccccctgctcacc 1253
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    420 AGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTCACC 479

Qy   1254 agtttccgctgtaagaaggagggccccacccctcagtgtgcccatggttcagggtgaatgc 1313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 AGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAATGC 539

Qy   1314 ctcctcaagtaccagctccgtcccaggaggg 1344
          |||||||||||||||||||||||||||||
Db    540 CTCCTCAAGTACCAGCTCCGTCCCAGGAGGG 570


RESULT  13
AW510825/c
LOCUS       AW510825      499 bp    mRNA            EST       03-MAR-2000
DEFINITION  hd40b11.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2911965 3', mRNA sequence.
ACCESSION   AW510825
VERSION     AW510825.1  GI:7148903
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 499)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
```

```
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 470.
FEATURES             Location/Qualifiers
     source          1. .499
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2911965"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo. "
BASE COUNT       94 a    142 c    126 g    137 t
ORIGIN


  Query Match             15.1%;  Score 448;  DB 91;  Length 499;
  Best Local Similarity   99.8%;  Pred. No. 1e-225;
  Matches  498;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   2460 gaggatggggagcctcagcagaagcgggcccacacagaggagccacaggccaagaaggtc 2519
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    499 GAGGATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGCCAAGAAGGTC 440

Qy   2520 agagcccagtgaagatctgggagaccctgaactcagaaggctgtgtgtcttctgccccac 2579
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    439 AGAGCCCAGTGAAGATCTGGGAGACCCTGAACTCAGAAGGCTGTGTGTCTTCTGCCCCAC 380

Qy   2580 gcacgcacccgtatctgccctccttgctggtagaagctgaagagcacggtcccccaggag 2639
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    379 GCACGCACCCGTATCTGCCCTCCTTGCTGGTAGAAGCTGAAGAGCACGGTCCCCCAGGAG 320

Qy   2640 gcagctcaggataggtggtatggagctgtgccgaggcttgggctcccacataagcactag 2699
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    319 GCAGCTCAGGATAGGTGGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAG 260

Qy   2700 tctatagatgcctcttaggactggtgcctggcacagccgcgggccaggaggctgccacac 2759
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    259 TCTATAGATGCCTCTTAGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACAC 200

Qy   2760 ggaagcaagcagatgaactaatttcatttcaaggcagtttttaaagaagtcttggaaaca 2819
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    199 GGAAGCAAGCAGATGAACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCATGGAAACA 140

Qy   2820 gacggcggcacctttcctctaatccagcaaagtgattccctgcacaccagagacaagcag 2879
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    139 GACGGCGGCACCTTTCCTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAG 80

Qy   2880 agtaacaggatcagtgggtctaagtgtccgagacttaacgaaaatagtatttcagctgca 2939
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     79 AGTAACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCA 20

Qy   2940 ataaagattgagtttgcaa 2958
           ||||||||||||||||||
Db     19 ATAAAGATTGAGTTTGCAA 1


RESULT  14
```

```
BE260626
LOCUS       BE260626      938 bp    mRNA            EST       13-JUL-2000
DEFINITION  601146116F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3161691 5',
            mRNA sequence.
ACCESSION   BE260626
VERSION     BE260626.1  GI:9132065
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 938)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM119  row: o  column: 04
            High quality sequence stop: 621.
FEATURES             Location/Qualifiers
     source          1. .938
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3161691"
                     /clone_lib="NIH_MGC_19"
                     /tissue_type="neuroblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and  Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
BASE COUNT      190 a    263 c    283 g    202 t
ORIGIN


  Query Match             15.1%;  Score 447;  DB 106;  Length 938;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-225;
  Matches  447;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    966 tttgtggtggtagaatgtccagatgaaagcttcattcaacccatctgtgagaatgccacc 1025
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    135 TTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATCTGTGAGAATGCCACC 194

Qy   1026 tttcagaggtaccaaggaaaggcagatgcccccgtggccttggtggttcacatggcccca 1085
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    195 TTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTGGTTCACATGGCCCCA 254

Qy   1086 gcatctgtgcttgtggacagcaggtaccagcagtggatggagaggtttgggcctgacacc 1145
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    255 GCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGGTTTGGGCCTGACACC 314

Qy   1146 cagcacttggtcctgaatgagaactgtgcctcagttcacaaccttcgcagccacaagatt 1205
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 CAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTTCGCAGCCACAAGATT 374

Qy   1206 caaacccagctcaacctcatccacccggacatcttccccctgctcaccagtttccgctgt 1265
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 CAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTCACCAGTTTCCGCTGT 434

Qy   1266 aagaaggagggccccacccctcagtgtgcccatggttcagggtgaatgcctcctcaagtac 1325
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    435 AAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAATGCCTCCTCAAGTAC 494
```

```
Qy   1326 cagctccgtcccaggagggagtggcagagggatgccattattacttgcaatcctgaggaa 1385
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    495 CAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACTTGCAATCCTGAGGAA 554

Qy   1386 ttcatagttgaggcgctgcagcttccc 1412
          |||||||||||||||||||||||||||
Db    555 TTCATAGTTGAGGCGCTGCAGCTTCCC 581


RESULT  15
AW575677
LOCUS       AW575677      493 bp    mRNA            EST       15-MAR-2000
DEFINITION  UI-HF-BM0-adl-b-07-0-UI.s1 NIH_MGC_38 Homo sapiens cDNA clone
            IMAGE:3061957 3', mRNA sequence.
ACCESSION   AW575677
VERSION     AW575677.1  GI:7247216
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 493)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message. Tissue Procurement: Louis M.
            Staudt, M.D., Ph.D.
             cDNA Library Preparation: M.B. Soares Lab
             cDNA Library Arrayed by: M.B. Soares Lab
             DNA Sequencing by: M.B. Soares Lab
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: M13 Forward
            POLYA=No.
FEATURES              Location/Qualifiers
     source           1. .493
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:3061957"
                      /clone_lib="NIH_MGC_38"
                      /tissue_type="lymph"
                      /cell_type="germinal center B cells"
                      /cell_line="MGC85"
                      /lab_host="DH10B (LTI)"
                      /note="Vector: pT7T3-Pac; Site_1: NotI; Site_2: Eco RI;
                      Constructed from size fractionated cytoplasmic mRNA
                      (2.5-3.5kb). Directionally cloned. Cells provided by Louis
                      M. Staudt, Ph.D. Library preparation by Maria de Fatima
                      Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
BASE COUNT      126 a    131 c    138 g     97 t      1 others
ORIGIN

  Query Match           14.9%;  Score 442;  DB 92;  Length 493;
  Best Local Similarity   99.8%;  Pred. No. 1.5e-222;
  Matches  492;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    141 cggccgcgcaaggacccgctgcggcacctgcgcacgcgagagaagcgcggaccgtcgggg 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CGGCCGCGCAAGGACCCGCTGCGGCACCTGCGCACGCGAGAGAAGCGCGGACCGTCGGGG 60

Qy    201 tgctccggcggcccaaacaccgtgtacctgcaggtggtggcagcgggtagccgggactcg 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGCTCCGGCGGCCCAAACACCGTGTACCTGCAGGTGGTGGCAGCGGGTAGCCGGGACTCG 120

Qy    261 ggcgccgcgctctacgtcttctccgagttcaaccggtatctcttcaactgtggagaaggc 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    121 GGCGCCGCGCTCTACGTCTTCTCCGAGTTCAACCGGTATCTCTTCAACTGTGGAGAAGGC 180

Qy    321 gttcagagactcatgcaggagcacaagttaaaggttgctcgcctggacaacatattcctg 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTTCAGAGACTCATGCAGGAGCACAAGTTAAAGGTTGCTCGCCTGGACAACATATTCCTG 240

Qy    381 acacgaatgcactggtctaatgttgggggcttaagtggaatgattcttactttaaaggaa 440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ACACGAATGCACTGGTCTAATGTTGGGGGCTTAAGTGGAATGATTCTTACTTTAAAGGAA 300

Qy    441 accgggcttccaaagtgtgtactttctggacctccacaactggaaaaatacctcgaagca 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ACCGGGCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATACCTCGAAGCA 360

Qy    501 atcaaaatattttctggtccattgaaaggaatagaactggctgtgcggccccactctgcc 560
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    361 ATCAAAATATTTTCTGGTCCATTGAAAGGAATAGAACTGGCTGTGCGGNCCCACTCTGCC 420

Qy    561 ccagaatacgaggatgaaaccatgacagtttaccagatccccatacacagtgaacagagg 620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCAGAATACGAGGATGAAACCATGACAGTTTACCAGATCCCCATACACAGTGAACAGAGG 480

Qy    621 aggggaaagcacc 633
          |||||||||||||
Db    481 AGGGGAAAGCACC 493


RESULT  16
BE304720
LOCUS       BE304720      612 bp    mRNA            EST       13-JUL-2000
DEFINITION  601106236F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:3349304 5',
            mRNA sequence.
ACCESSION   BE304720
VERSION     BE304720.1  GI:9176150
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 612)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: NIH Intramural Sequencing Center
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM143  row: h  column: 09
            High quality sequence start: 21
            High quality sequence stop: 609.
FEATURES              Location/Qualifiers
     source           1. .612
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:3349304"
                      /clone_lib="NIH_MGC_15"
                      /tissue_type="adenocarcinoma cell line"
                      /lab_host="DH10B (phage-resistant)"
                      /note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
                      EcoRI; cDNA made by oligo-dT priming.  Directionally
                      cloned into EcoRI/XhoI sites using the following 5'
                      adaptor: GGCACGAG(G).  Size-selected >500bp for average
                      insert size 1.8kb.  Library constructed by Ling Hong in
                      the laboratory of Gerald M. Rubin (University of
                      California, Berkeley) using  ZAP-cDNA synthesis kit
                      (Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT      143 a    178 c    161 g    130 t
ORIGIN
```

```
  Query Match             14.8%;  Score 438;  DB 106;  Length 612;
  Best Local Similarity   100.0%;  Pred. No. 2e-220;
  Matches  438;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1041 ggaaaggcagatgcccccgtggccttggtggttcacatggccccagcatctgtgcttgtg 1100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     108 GGAAAGGCAGATGCCCCCGTGGCCTTGGTGGTTCACATGGCCCCAGCATCTGTGCTTGTG 167

Qy    1101 gacagcaggtaccagcagtggatggagaggtttgggcctgacacccagcacttggtcctg 1160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168 GACAGCAGGTACCAGCAGTGGATGGAGAGGTTTGGGCCTGACACCCAGCACTTGGTCCTG 227

Qy    1161 aatgagaactgtgcctcagttcacaaccttcgcagccacaagattcaaacccagctcaac 1220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     228 AATGAGAACTGTGCCTCAGTTCACAACCTTCGCAGCCACAAGATTCAAACCCAGCTCAAC 287

Qy    1221 ctcatccacccggacatcttccccctgctcaccagtttccgctgtaagaaggagggcccc 1280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     288 CTCATCCACCCGGACATCTTCCCCCTGCTCACCAGTTTCCGCTGTAAGAAGGAGGGCCCC 347

Qy    1281 accctcagtgtgcccatggttcagggtgaatgcctcctcaagtaccagctccgtcccagg 1340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     348 ACCCTCAGTGTGCCCATGGTTCAGGGTGAATGCCTCCTCAAGTACCAGCTCCGTCCCAGG 407

Qy    1341 agggagtggcagagggatgccattattacttgcaatcctgaggaattcatagttgaggcg 1400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     408 AGGGAGTGGCAGAGGGATGCCATTATTACTTGCAATCCTGAGGAATTCATAGTTGAGGCG 467

Qy    1401 ctgcagcttcccaacttccagcagagcgtgcaggagtacaggaggagtgcgcaggacggc 1460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     468 CTGCAGCTTCCCAACTTCCAGCAGAGCGTGCAGGAGTACAGGAGGAGTGCGCAGGACGGC 527

Qy    1461 ccagccccagcagagaaa 1478
           ||||||||||||||||||
Db     528 CCAGCCCCAGCAGAGAAA 545


RESULT  17
AA455121
LOCUS       AA455121      493 bp    mRNA            EST       06-JUN-1997
DEFINITION  zx78c04.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
            IMAGE:809862 3' similar to SW:YK59_YEAST P36159 HYPOTHETICAL 96.8
            KD PROTEIN IN SIS2-MTD1 INTERGENIC REGION. ;, mRNA sequence.
ACCESSION   AA455121
VERSION     AA455121.1  GI:2177897
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 493)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
            , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
  TITLE     WashU-Merck EST Project 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: similarity on wrong strand
            Possible reversed clone: polyT not found
            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 447.
FEATURES             Location/Qualifiers
     source          1. .493
                     /organism="Homo sapiens"
                     /db_xref="GDB:6039680"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:809862"
                     /clone_lib="Soares ovary tumor NbHOT"
                     /sex="Female"
                     /tissue_type="ovarian tumor"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo. "
BASE COUNT      131 a    128 c    130 g    104 t
ORIGIN


  Query Match             14.7%;  Score 436;  DB 7;  Length 493;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-219;
  Matches  436;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     252 cgggactcgggcgccgcgctctacgtcttctccgagttcaaccggtatctcttcaactgt 311
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      58 CGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTCAACCGGTATCTCTTCAACTGT 117

Qy     312 ggagaaggcgttcagagactcatgcaggagcacaagttaaaggttgctcgcctggacaac 371
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     118 GGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTAAAGGTTGCTCGCCTGGACAAC 177

Qy     372 atattcctgacacgaatgcactggtctaatgttgggggcttaagtggaatgattcttact 431
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     178 ATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGCTTAAGTGGAATGATTCTTACT 237

Qy     432 ttaaaggaaaccgggcttccaaagtgtgtactttctggacctccacaactggaaaaatac 491
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     238 TTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATAC 297

Qy     492 ctcgaagcaatcaaaatattttctggtccattgaaaggaatagaactggctgtgcggccc 551
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     298 CTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGAATAGAACTGGCTGTGCGGCCC 357

Qy     552 cactctgccccagaatacgaggatgaaaccatgacagtttaccagatccccatacacagt 611
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     358 CACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTTTACCAGATCCCCATACACAGT 417

Qy     612 gaacagaggaggggaaagcaccaaccatggcagagtccagaaaggcctctcagcaggctc 671
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     418 GAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGTCCAGAAAGGCCTCTCAGCAGGCTC 477

Qy     672 agtccagagcgatctt 687
           ||||||||||||||||
Db     478 AGTCGAGAGCGATCTT 493


RESULT  18
AW592601/c
LOCUS       AW592601      478 bp    mRNA            EST       22-MAR-2000
DEFINITION  hf45a09.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2934808 3', mRNA sequence.
ACCESSION   AW592601
VERSION     AW592601.1  GI:7279786
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 478)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
```

```
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
OMMENT      Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 451.
EATURES             Location/Qualifiers
     source         1. .478
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:2934808"
                    /clone_lib="Soares_NFL_T_GBC_S1"
                    /lab_host="DH10B"
                    /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                    a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                    Equal amounts of plasmid DNA from three normalized
                    libraries (fetal lung NbHL19W, testis NHT, and B-cell
                    NCI_CGAP_GCB1) were mixed, and ss circles were made in
                    vitro. Following HAP purification, this DNA was used as
                    tracer in a subtractive hybridization reaction. The driver
                    was PCR-amplified cDNAs from pools of 5,000 clones made
                    from the same 3 libraries. The pools consisted of
                    I.M.A.G.E. clones 297480-302087, 682632-687239,
                    726408-728711, and 729096-731399. Subtraction by Bento
                    Soares and M. Fatima Bonaldo. "
ASE COUNT       92 a    132 c    122 g    132 t
RIGIN
```

```
  Query Match             14.4%;  Score 425;  DB 92;  Length 478;
  Best Local Similarity   99.8%;  Pred. No. 1.5e-213;
  Matches  475;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   2483 gcgggcccacacagaggagccacaggccaagaaggtcagagcccagtgaagatctgggag 2542
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    478 GCGGGCCCACACAGAGGAGCCACAGGCCAAGAAGGTCAGAGCCCAGTGAAGATCTGGGAG 419

Qy   2543 accctgaactcagaaggctgtgtgtcttctgccccacgcacgcacccgtatctgccctcc 2602
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    418 ACCCTGAACTCAGAAGGCTGTGTGTCTTCTGCCCCACGCACGCACCCGTATCTGCCCTCC 359

Qy   2603 ttgctggtagaagctgaagagcacggtcccccaggaggcagctcaggataggtggtatgg 2662
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    358 TTGCTGGTAGAAGCTGAAGAGCACGGTCCCCCAGGAGGCAGCTCAGGATAGGTGGTATGG 299

Qy   2663 agctgtgccgaggcttgggctcccacataagcactagtctatagatgcctcttaggactg 2722
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    298 AGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTTAGGACTG 239

Qy   2723 gtgcctggcacagccgcgggccaggaggctgccacacggaagcaagcagatgaactaatt 2782
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    238 GTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGAACTAATT 179

Qy   2783 tcatttcaaggcagtttttaaagaagtcttggaaacagacggcggcacctttcctctaat 2842
           |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    178 TCATTTCAAGGCAGTTTTTAAAGAAGTCATGGAAACAGACGGCGGCACCTTTCCTCTAAT 119

Qy   2843 ccagcaaagtgattccctgcacaccagagacaagcagagtaacaggatcagtgggtctaa 2902
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    118 CCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTGGGTCTAA 59

Qy   2903 gtgtccgagacttaacgaaaatagtatttcagctgcaataaagattgagtttgcaa 2958
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     58 GTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTGCAA 3
```

```
ESULT  19
I089646/c
OCUS       AI089646      527 bp    mRNA            EST       18-AUG-1998
EFINITION  qb16g07.x1 Soares_pregnant_uterus_NbHPU Homo sapiens cDNA clone
           IMAGE:1696476 3', mRNA sequence.
ACCESSION  AI089646
VERSION    AI089646.1  GI:3428705
KEYWORDS   EST.
SOURCE     human.
  ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 527)
  AUTHORS  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE    National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
           Tumor Gene Index
  JOURNAL  Unpublished (1997)
COMMENT    Contact: Robert Strausberg, Ph.D.
           Tel: (301) 496-1550
           Email: Robert_Strausberg@nih.gov
           This clone is available royalty-free through LLNL ; contact the
           IMAGE Consortium (info@image.llnl.gov) for further information.
           Seq primer: -40m13 fwd. ET from Amersham
           High quality sequence stop: 463.
FEATURES            Location/Qualifiers
     source         1. .527
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:1696476"
                    /clone_lib="Soares_pregnant_uterus_NbHPU"
                    /sex="female"
                    /dev_stage="adult"
                    /lab_host="DH10B"
                    /note="Organ: uterus; Vector: pT7T3-Pac; Site_1: Not I;
                    Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
                    oligo(dT) primer [5'
                    AACTGGAAGAATTCGCGGCCGCCTTTTTTTTTTTTTTTT 3'],
                    double-stranded cDNA was ligated to Eco RI adaptors
                    (Pharmacia), digested with Not I and cloned into the Not I
                    and Eco RI sites of the modified pT7T3 vector. Library
                    went through one round of normalization.  Library
                    constructed by M. Fatima Bonaldo."
BASE COUNT      99 a    153 c    136 g    138 t      1 others
ORIGIN
```

```
  Query Match             14.4%;  Score 425;  DB 16;  Length 527;
  Best Local Similarity   99.6%;  Pred. No. 1.5e-213;
  Matches  525;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy   2430 ctcctgtccagggagctggcaggcggcctggaggatggggagcctcagcagaagcgggcc 2489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    527 CTCCTGTCCAGGGAGCTGGCAGGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCC 468

Qy   2490 cacacagaggagccacaggccaagaaggtcagagcccagtgaagatctgggagaccctga 2549
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    467 CACACAGAGGAGCCACAGGCCAAGAAGGTCAGAGCCCAGTGAAGATCTGGGAGACCCTGA 408

Qy   2550 actcagaaggctgtgtgtcttctgccccacgcacgcacccgtatctgccctccttgctgg 2609
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    407 ACTCAGAAGNCTGTGTGTCTTCTGCCCCACGCACGCACCCGTATCTGCCCTCCTTGCTGG 348

Qy   2610 tagaagctgaagagcacggtcccccaggaggcagctcaggataggtggtatggagctgtg 2669
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    347 TAGAAGCTGAAGAGCACGGTCCCCCAGGAGGCAGCTCAGGATAGGTGGTATGGAGCTGTG 288

Qy   2670 ccgaggcttgggctcccacataagcactagtctatagatgcctcttaggactggtgcctg 2729
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    287 CCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTTAGGACTGGTGCCTG 228

Qy   2730 gcacagccgcgggccaggaggctgccacacggaagcaagcagatgaactaatttcatttc 2789
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    227 GCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGAACTAATTTCATTTC 168

Qy   2790 aaggcagtttttaaagaagtcttggaaacagacggcggcacctttcctctaatccagcaa 2849
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    167 AAGGCAGTTTTTAAAGAAGTCATGGAAACAGACGGCGGCACCTTTCCTCTAATCCAGCAA 108
```

```
Qy   2850 agtgattccctgcacaccagagacaagcagagtaacaggatcagtgggtctaagtgtccg 2909
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    107 AGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTGGGTCTAAGTGTCCG 48

Qy   2910 agacttaacgaaaatagtatttcagctgcaataaagattgagtttgc 2956
          |||||||||||||||||||||||||||||||||||||||||||||
Db     47 AGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTGC 1


RESULT  20
AI468143
LOCUS       AI468143      421 bp    mRNA            EST       30-MAR-1999
DEFINITION  tf92g05.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:2106776 3'
            similar to SW:YK59_YEAST P36159 HYPOTHETICAL 96.8 KD PROTEIN IN
            SIS2-MTD1 INTERGENIC REGION. ;, mRNA sequence.
ACCESSION   AI468143
VERSION     AI468143.1  GI:4330233
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 421)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
            Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 733   Std Error: 0.00
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .421
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2106776"
                     /clone_lib="NCI_CGAP_Brn23"
                     /tissue_type="glioblastoma (pooled)"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATATCTTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT       98 a    114 c    121 g     88 t
ORIGIN


  Query Match          14.2%;  Score 421;  DB 20;  Length 421;
  Best Local Similarity   100.0%;  Pred. No. 2e-211;
  Matches  421;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    140 gcggccgcgcaaggacccgctgcggcacctgcgcacgcgagagaagcgcggaccgtcggg 199
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCGGCCGCGCAAGGACCCGCTGCGGCACCTGCGCACGCGAGAGAAGCGCGGACCGTCGGG 60

Qy    200 gtgctccggcggcccaaacaccgtgtacctgcaggtggtggcagcgggtagccgggactc 259
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GTGCTCCGGCGGCCCAAACACCGTGTACCTGCAGGTGGTGGCAGCGGGTAGCCGGGACTC 120

Qy    260 gggcgccgcgctctacgtcttctccgagttcaaccggtatctcttcaactgtggagaagg 319
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GGGCGCCGCGCTCTACGTCTTCTCCGAGTTCAACCGGTATCTCTTCAACTGTGGAGAAGG 180

Qy    320 cgttcagagactcatgcaggagcacaagttaaaggttgctcgcctggacaacatattcct 379
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CGTTCAGAGACTCATGCAGGAGCACAAGTTAAAGGTTGCTCGCCTGGACAACATATTCCT 240

Qy    380 gacacgaatgcactggtctaatgttgggggcttaagtggaatgattcttactttaaagga 439
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GACACGAATGCACTGGTCTAATGTTGGGGGCTTAAGTGGAATGATTCTTACTTTAAAGGA 300

Qy    440 aaccgggcttccaaagtgtgtactttctggacctccacaactggaaaaatacctcgaagc 499
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AACCGGGCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATACCTCGAAGC 360

Qy    500 aatcaaaatatttttctggtccattgaaaggaatagaactggctgtgcggccccactctgc 559
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AATCAAAATATTTTCTGGTCCATTGAAAGGAATAGAACTGGCTGTGCGGCCCCACTCTGC 420

Qy    560 c 560
          |
Db    421 C 421


RESULT  21
AW206103
LOCUS       AW206103      536 bp    mRNA            EST       02-DEC-1999
DEFINITION  UI-H-BI1-afx-f-12-0-UI.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone
            IMAGE:2723471 3', mRNA sequence.
ACCESSION   AW206103
VERSION     AW206103.1  GI:6505577
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 536)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message. cDNA Library Preparation: M.B.
            Soares Lab Clone distribution: NCI-CGAP clone distribution
            information can be found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: M13 Forward
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2723471"
                     /clone_lib="NCI_CGAP_Sub3"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NCI_CGAP_Sub3 library is a subtracted library derived from
                     the NCI_CGAP_Sub1 library, which is a subtracted library
                     derived from BI. BI constitutes a mixture of 21
                     normalized or subtracted NCI_CGAP libraries: NCI_CGAP_Co4
                     , NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_Co10,
                     NCI_CGAP_Co16, NCI_CGAP_Kid5, NCI_CGAP_Kid12,
                     NCI_CGAP_Kid3, NCI_CGAP_Kid11, NCI_CGAP_Lym2,
                     NCI_CGAP_Br2, NCI_CGAP_Co8, NCI_CGAP_CLL1, NCI_CGAP_Lei2,
                     NCI_CGAP_Brn23, NCI_CGAP_Lu5, NCI_CGAP_Lu24,
```

NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC6, NCI_CGAP_Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below: NCI_CGAP_Kid3 pool 1 LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE CloneIDs 1322376-1323911, 1456008-1456775, 1500552-1502855); NCI_CGAP_Kid5 pool 1 LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE CloneIDs 1323912-1325831, 1471368-1472903, 1492104-1493255); NCI_CGAP_Lu5 pool 1 LLAM 3575-3582, 3851-3854 (IMAGE CloneIDs 1414920-1417991, 1520904-1522439); NCI_CGAP_GC4 pool 1 LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE CloneIDs 1257096-1258631, 1469064-1470983, 1475592-1476743); NCI_CGAP_Pr22 pool 1 LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE CloneIDs 985608-986759, 1101192-1101959, 1217928-1220615); NCI_CGAP_Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE CloneIDs 1057416-1061255, 1144584-1145351). Subtraction was performed as previously described [Bonaldo, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806.
TAG_LIB=NCI_CGAP_Lu5
TAG_TISSUE=lung
TAG_SEQ=CAAC"

ASE COUNT 129 a 143 c 146 g 118 t
RIGIN

Query Match 14.2%; Score 421; DB 87; Length 536;
Best Local Similarity 99.8%; Pred. No. 2e-211;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
y   141 cggccgcgcaaggacccgctgcggcacctgcgcacgcgagagaagcgcggaccgtcgggg 200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
    1 CGGCCGCGCAAGGACCCGCTGCGGCACCTGCGCACGCGAGAGAAGCGCGGACCGTCGGGG 60

y   201 tgctccggcggcccaaacaccgtgtacctgcaggtggtggcagcgggtagccgggactcg 260
        |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
   61 TGCTCCGGCGGCCCAAACACCGTGTACCTGCAGGTGGTTGCAGCGGGTAGCCGGGACTCG 120

y   261 ggcgccgcgctctacgtcttctccgagttcaaccggtatctcttcaactgtggagaaggc 320
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   121 GGCGCCGCGCTCTACGTCTTCTCCGAGTTCAACCGGTATCTCTTCAACTGTGGAGAAGGC 180

y   321 gttcagagactcatgcaggagcacaagttaaaggttgctcgcctggacaacatattcctg 380
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   181 GTTCAGAGACTCATGCAGGAGCACAAGTTAAAGGTTGCTCGCCTGGACAACATATTCCTG 240

y   381 acacgaatgcactggtctaatgttgggggcttaagtggaatgattcttactttaaaggaa 440
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   241 ACACGAATGCACTGGTCTAATGTTGGGGGCTTAAGTGGAATGATTCTTACTTTAAAGGAA 300

y   441 accgggcttccaaagtgtgtactttctggacctccacaactggaaaaatacctcgaagca 500
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   301 ACCGGGCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATACCTCGAAGCA 360

y   501 atcaaaatattttctggtccattgaaaggaatagaactggctgtgcggcctcactctgcc 560
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   361 ATCAAAATATTTTCTGGTCCATTGAAAGGAATAGAACTGGCTGTGCGGCCCCACTCTGCC 420

y   561 ccagaatacgaggatgaaaccatgacagtttaccagatccccatacacagtg 612
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
   421 CCAGAATACGAGGATGAAACCATGACAGTTTACCAGATCCCCATACACAGTG 472
```

ESULT 22
W304130
OCUS AW304130 404 bp mRNA EST 18-JAN-2000
EFINITION xs13e05.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2769536 3' similar to SW:YK59_YEAST P36159 HYPOTHETICAL 96.8 KD PROTEIN IN SIS2-MTD1 INTERGENIC REGION. ;, mRNA sequence.
CCESSION AW304130
VERSION AW304130.1 GI:6713819
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 404)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Possible reversed clone: similarity on wrong strand
Possible reversed clone: polyT not found
Seq primer: -40UP from Gibco.
FEATURES Location/Qualifiers
source 1. .404
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2769536"
/clone_lib="NCI_CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 96 a 107 c 113 g 88 t
ORIGIN

Query Match 13.7%; Score 404; DB 88; Length 404;
Best Local Similarity 100.0%; Pred. No. 2e-202;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy  160 tgcggcacctgcgcacgcgagagaagcgcggaccgtcggggtgctccggcggcccaaaca 219
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 TGCGGCACCTGCGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACA 60

Qy  220 ccgtgtacctgcaggtggtggcagcgggtagccgggactcgggcgccgcgctctacgtct 279
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CCGTGTACCTGCAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCT 120

Qy  280 tctccgagttcaaccggtatctcttcaactgtggagaaggcgttcagagactcatgcagg 339
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 TCTCCGAGTTCAACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGG 180

Qy  340 agcacaagttaaaggttgctcgcctggacaacatattcctgacacgaatgcactggtcta 399
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AGCACAAGTTAAAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTA 240

Qy  400 atgttgggggcttaagtggaatgattcttactttaaaggaaaccgggcttccaaagtgtg 459
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 ATGTTGGGGGCTTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTG 300

Qy  460 tactttctggacctccacaactggaaaaatacctcgaagcaatcaaaatattttctggtc 519
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    301 TACTTTCTGGACCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTC 360

Qy    520 cattgaaaggaatagaactggctgtgcggccccactctgcccca 563
          |||||||||||||||||||||||||||||||||||||||||||||
Db    361 CATTGAAAGGAATAGAACTGGCTGTGCGGCCCCACTCTGCCCCA 404


RESULT  23
AA291670
LOCUS       AA291670      474 bp    mRNA            EST       16-MAY-1997
DEFINITION  zt37d04.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
            IMAGE:724519 3' similar to SW:YK59_YEAST P36159 HYPOTHETICAL 96.8
            KD PROTEIN IN SIS2-MTD1 INTERGENIC REGION. ;, mRNA sequence.
ACCESSION   AA291670
VERSION     AA291670.1  GI:1939648
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 474)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
            , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
  TITLE     WashU-Merck EST Project 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: similarity on wrong strand
            Possible reversed clone: polyT not found
            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 465.
FEATURES             Location/Qualifiers
     source          1. .474
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:724519"
                     /clone_lib="Soares ovary tumor NbHOT"
                     /sex="Female"
                     /tissue_type="ovarian tumor"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo. "
BASE COUNT      128 a    120 c    129 g     97 t
ORIGIN


  Query Match          13.6%;  Score 402;  DB 4;  Length 474;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-201;
  Matches  402;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    251 ccgggactcgggcgccgcgctctacgtcttctccgagttcaaccggtatctcttcaactg 310
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     73 CCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTCAACCGGTATCTCTTCAACTG 132

Qy    311 tggagaaggcgttcagagactcatgcaggagcacaagttaaaggttgctcgcctggacaa 370
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    133 TGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTAAAGGTTGCTCGCCTGGACAA 192

Qy    371 catattcctgacacgaatgcactggtctaatgttgggggcttaagtggaatgattcttac 430
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    193 CATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGCTTAAGTGGAATGATTCTTAC 252

Qy    431 tttaaaggaaaccgggcttccaaagtgtgtactttctggacctccacaactggaaaaata 490
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    253 TTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATA 312

Qy    491 cctcgaagcaatcaaaatattttctggtccattgaaaggaatagaactggctgtgcggcc 550
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    313 CCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGAATAGAACTGGCTGTGCGGCC 372

Qy    551 ccactctgccccagaatacgaggatgaaaccatgacagtttaccagatccccatacacag 610
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    373 CCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTTTACCAGATCCCCATACACAG 432

Qy    611 tgaacagaggaggggaaagcaccaaccatggcagagtccaga 652
          |||||||||||||||||||||||||||||||||||||||||
Db    433 TGAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGTCCAGA 474


RESULT  24
BE747163
LOCUS       BE747163      992 bp    mRNA            EST       15-SEP-2000
DEFINITION  601577254F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3838237 5',
            mRNA sequence.
ACCESSION   BE747163
VERSION     BE747163.1  GI:10161155
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 992)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM524  row: l  column: 14
            High quality sequence stop: 781.
FEATURES             Location/Qualifiers
     source          1. .992
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3838237"
                     /clone_lib="NIH_MGC_9"
                     /tissue_type="adenocarcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      225 a    287 c    313 g    167 t
ORIGIN


  Query Match          13.5%;  Score 400;  DB 135;  Length 992;
  Best Local Similarity   99.7%;  Pred. No. 2.8e-200;
  Matches  570;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy   1132 ttgggcctgacacccagcacttggtcctgaatgagaactgtgcctcagttcacaaccttc 1191
```

```
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       90 TTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTTC 149

Qy     1192 gcagccacaagattcaaacccagctcaacctcatccacccggacatcttcccсctgctca 1251
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      150 GCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTCA 209

Qy     1252 ccagtttccgctgtaagaaggagggccccaccctcagtgtgcccatggttcagggtgaat 1311
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      210 CCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAAT 269

Qy     1312 gcctcctcaagtaccagctccgtcccaggagggagtggcagagggatgccattattactt 1371
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      270 GCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACTT 329

Qy     1372 gcaatcctgaggaattcatagttgaggcgctgcagcttcccaacttccagcagagcgtgc 1431
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      330 GCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTGC 389

Qy     1432 aggagtacaggaggagtgcgcaggacggcccagccccagcagagaaaagaagtcagtacc 1491
            |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db      390 AGGAGTACAGGAGGAGTGCGCAGGACGG-CCAGCCCCAGCAGAGAAAAGAAGTCAGTACC 448

Qy     1492 cagaaatcatcttccttggaacagggtctgccatcccgatgaagattcgaaatgtcagtg 1551
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      449 CAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGTG 508

Qy     1552 ccacacttgtcaacataagccccgacacgtctctgctactggactgtggtgagggcacat 1611
            |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db      509 CCACACTTGTCAACATAAGCCCCGACACGTCTTTGCTACTGGACTGTGGTGAGGGCACAT 568

Qy     1612 ttgggcagctgtgccgtcattacggagaccaggtggacagggtcctgggcaccctggctg 1671
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      569 TTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCTG 628

Qy     1672 ctgtgtttgtgtcccacctgcacgcagatcac 1703
            ||||||||||||||||||||||||||||||||
Db      629 CTGTGTTTGTGTCCCACCTGCACGCAGATCAC 660
```

```
RESULT 25
BE250309
LOCUS       BE250309     574 bp    mRNA            EST       13-JUL-2000
DEFINITION  600943455F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2960077 5',
            mRNA sequence.
ACCESSION   BE250309
VERSION     BE250309.1  GI:9120418
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 574)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM52  row: n  column: 14
            High quality sequence stop: 571.
FEATURES             Location/Qualifiers
     source          1..574
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2960077"
                     /clone_lib="NIH_MGC_17"
                     /tissue_type="rhabdomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
                     Site_2: XhoI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Size-selected >500bp
                     for average insert size 1.8kb.  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      124 a    158 c    167 g    119 t      6 others
ORIGIN

  Query Match             13.0%;  Score 384;  DB 105;  Length 574;
  Best Local Similarity   99.8%;  Pred. No. 8.2e-192;
  Matches  434;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     1624 gccgtcattacggagaccaggtggacagggtcctgggcaccctggctgctgtgtttgtgt 1683
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 GCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCTGCTGTGTTTGTGT 60

Qy     1684 cccacctgcacgcagatcaccacacgggcttgccaagtatcttgctgcagagagaacgcg 1743
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 CCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTGCAGAGAGAACGCG 120

Qy     1744 ccttggcatctttgggaaagccgcttcaccctttgctggtggttgcccccaaccagctca 1803
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 CCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCCCCCAACCAGCTCA 180

Qy     1804 aagcctggctccagcagtaccacaaccagtgccaggaggtcctgcaccacatcagtatga 1863
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 AAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCACCACATCAGTATGA 240

Qy     1864 ttcctgccaaatgccttcaggaaggggctgagatctccagtcctgcagtggaaagattga 1923
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 TTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCAGTGGAAAGATTGA 300

Qy     1924 tcagttcgctgttgcgaacatgtgatttggaagagtttcagacctgtctggtgcggcact 1983
            ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      301 TCAGTTCGCTGTTGCGAACGTGTGATTTGGAAGAGTTTCAGACCTGTCTGGTGCGGCACT 360

Qy     1984 gcaagcatgcgtttggctgtgcgctggtgcacacctctggctggaaagtggtctattccg 2043
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 GCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAAGTGGTCTATTCCG 420

Qy     2044 gggacaccatgccct 2058
            |||||||||||||||
Db      421 GGGACACCATGCCCT 435
```

```
RESULT 26
AA310236
LOCUS       AA310236     431 bp    mRNA            EST       19-APR-1997
DEFINITION  EST181085 Jurkat T-cells V Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION   AA310236
VERSION     AA310236.1  GI:1962584
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 431)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
            ,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
            ,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
            Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
            ,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,
            Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M.,
            Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
```

```
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.
            , Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
            Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L.,
            Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
            ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
            Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT     Other_ESTs: THC175624
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES              Location/Qualifiers
     source           1. .431
                      /organism="Homo sapiens"
                      /db_xref="ATCC (inhost):156306"
                      /db_xref="taxon:9606"
                      /clone_lib="Jurkat T-cells V"
                      /cell_type="T-lymphocyte"
                      /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                      XhoI"
BASE COUNT       98 a    131 c    117 g     84 t      1 others
ORIGIN


  Query Match             12.8%;  Score 380;  DB 5;  Length 431;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-189;
  Matches  430;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   1048 cagatgccccccgtggccttggtggttcacatggccccagcatctgtgcttgtggacagca 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CAGATGCCCCCGTGGCCTTGGTGGTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCA 60

Qy   1108 ggtaccagcagtggatggagaggtttgggcctgacacccagcacttggtcctgaatgaga 1167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGTACCAGCAGTGGATGGAGAGGTTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGA 120

Qy   1168 actgtgcctcagttcacaaccttcgcagccacaagattcaaacccagctcaacctcatcc 1227
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ACTGTGCCTCANTTCACAACCTTCGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCC 180

Qy   1228 acccggacatcttcccccctgctcaccagtttccgctgtaagaaggagggccccacccctca 1287
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCCGGACATCTTCCCCCTGCTCACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCA 240

Qy   1288 gtgtgcccatggttcagggtgaatgcctcctcaagtaccagctccgtcccaggagggagt 1347
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTGTGCCCATGGTTCAGGGTGAATGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGT 300

Qy   1348 ggcagagggatgccattattacttgcaatcctgaggaattcatagttgaggcgctgcagc 1407
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGCAGAGGGATGCCATTATTACTTGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGC 360

Qy   1408 ttcccaacttccagcagagcgtgcaggagtacaggaggagtgcgcaggacggcccagccc 1467
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TTCCCAACTTCCAGCAGAGCGTGCAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCC 420

Qy   1468 cagcagagaaa 1478
          ||||||||||||
Db    421 CAGCAGAGAAA 431
```

```
RESULT  27
BE744197
LOCUS       BE744197      949 bp    mRNA            EST       15-SEP-2000
DEFINITION  601577168F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3838121 5',
            mRNA sequence.
ACCESSION   BE744197
VERSION     BE744197.1  GI:10158189
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 949)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM524  row: g  column: 18
            High quality sequence stop: 668.
FEATURES              Location/Qualifiers
     source           1. .949
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:3838121"
                      /clone_lib="NIH_MGC_9"
                      /tissue_type="adenocarcinoma cell line"
                      /lab_host="DH10B (phage-resistant)"
                      /note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
                      EcoRI; cDNA made by oligo-dT priming.  Directionally
                      cloned into EcoRI/XhoI sites using the following 5'
                      adaptor: GGCACGAG(G).  Size-selected >500bp for average
                      insert size 1.8kb.  Library constructed by Ling Hong in
                      the laboratory of Gerald M. Rubin (University of
                      California, Berkeley) using  ZAP-cDNA synthesis kit
                      (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      225 a    263 c    281 g    179 t      1 others
ORIGIN


  Query Match             12.6%;  Score 373;  DB 135;  Length 949;
  Best Local Similarity   99.7%;  Pred. No. 5.7e-186;
  Matches  613;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    2;

Qy    722 tccacatggtgttagccagagaagaggggtcagggactcttccctggtcgtagctttcat 781
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TCCACATGGTGTTAGCCAGAGAAGAGGGGTCAGGGACTCTTCCCTGGTCGTAGCTTTCAT 60

Qy    782 ctgtaagcttcacttaaagagaggaaacttcttggtgctcaaagcaaaggagatgggcct 841
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGTAAGCTTCACTTAAAGAGAGGAAACTTCTTGGTGCTCAAAGCAAAGGAGATGGGCCT 120

Qy    842 cccagttgggacagctgccatcgctcccatcattgctgctgtcaaggacgggaaaagcat 901
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAGGACGGGAAAAGCAT 180

Qy    902 cactcatgaaggaagagagattttggctgaagagctgtgtactcctccagatcctggtgc 961
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    181 CACTCATGAAGGAAGAGAGA-TTTGGCTGAAGAGCTGTGTACTCCTCCAGATCCTGGTGC 239

Qy    962 tgcttttgtggtggtagaatgtccagatgaaagcttcattcaacccatctgtgagaatgc 1021
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 TGC-TTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATCTGTGAGAATGC 298

Qy   1022 caccttcagaggtaccaaggaaaggcagatgcccccgtggccttggtggttcacatggc 1081
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    299 CACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTGGTTCACATGGC 358

Qy   1082 cccagcatctgtgcttgtggacagcaggtaccagcagtggatggagaggtttgggcctga 1141
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    359 CCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGGTTTGGGCCTGA 418

Qy   1142 cacccagcacttggtcctgaatgagaactgtgcctcagttcacaaccttcgcagccacaa 1201
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    419 CACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTTCGCAGCCACAA 478

Qy   1202 gattcaaacccagctcaacctcatccacccggacatcttccccctgctcaccagtttccg 1261
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    479 GATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTCACCAGTTTCCG 538

Qy   1262 ctgtaagaaggagggccccacccctcagtgtgcccatggttcagggtgaatgcctcctcaa 1321
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    539 CTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAATGCCTCCTCAA 598

Qy   1322 gtaccagctccgtcc 1336
          ||||||||||||||| 
Db    599 GTACCAGCTCCGTCC 613


RESULT  28
BE883616
LOCUS       BE883616      370 bp    mRNA            EST       27-SEP-2000
DEFINITION  601508091F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3909527 5',
            mRNA sequence.
ACCESSION   BE883616
VERSION     BE883616.1  GI:10332392
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 370)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM710  row: f  column: 24
            High quality sequence stop: 370.
FEATURES             Location/Qualifiers
     source          1. .370
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3909527"
                     /clone_lib="NIH_MGC_71"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
BASE COUNT       94 a    100 c    106 g     70 t
ORIGIN


  Query Match           12.5%;  Score 370;  DB 137;  Length 370;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-184;
  Matches  370;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2032 tggtctattccggggacaccatgccctgcgaggctctggtccggatggggaaagatgcca 2091
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCCA 60

Qy   2092 ccctcctgatacatgaagccaccctggaagatggtttggaagaggaagcagtggaaaaga 2151
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTGGAAGAGGAAGCAGTGGAAAAGA 120

Qy   2152 cacacagcacaacgtcccaagccatcagcgtggggatgcggatgaacgcggagttcatta 2211
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCATTA 180

Qy   2212 tgctgaaccacttcagccagcgctatgccaaggtccccctcttcagccccaacttcagcg 2271
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAGCG 240

Qy   2272 agaaagtgggagttgcctttgaccacatgaaggtctgctttggagactttccaacaatgc 2331
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 AGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACAATGC 300

Qy   2332 ccaagctgattcccccactgaaagccctgtttgctggcgacatcgaggagatggaggagc 2391
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGGAGGAGC 360

Qy   2392 gcagggagaa 2401
          ||||||||||
Db     361 GCAGGGAGAA 370


RESULT  29
BE900936
LOCUS       BE900936      745 bp    mRNA            EST       29-SEP-2000
DEFINITION  601674206F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3957240 5',
            mRNA sequence.
ACCESSION   BE900936
VERSION     BE900936.1  GI:10389609
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 745)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM834  row: k  column: 01
            High quality sequence stop: 734.
FEATURES             Location/Qualifiers
     source          1. .745
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3957240"
                     /clone_lib="NIH_MGC_21"
                     /tissue_type="choriocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Size-selected >500bp
                     for average insert size 1.8kb.  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      165 a    206 c    215 g    159 t
ORIGIN
```

```
  Query Match             12.4%;  Score 368;  DB 137;  Length 745;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-183;
  Matches  368;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1678 ttgtgtcccacctgcacgcagatcaccacacgggcttgccaagtatcttgctgcagagag 1737
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104 TTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTGCAGAGAG 163

Qy   1738 aacgcgccttggcatctttgggaaagccgcttcacccttgctggtggttgcccccaaccc 1797
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    164 AACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCCCCCAACC 223

Qy   1798 agctcaaagcctggctccagcagtaccacaaccagtgccaggaggtcctgcaccacatca 1857
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    224 AGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCACCACATCA 283

Qy   1858 gtatgattcctgccaaatgccttcaggaaggggctgagatctccagtcctgcagtggaaa 1917
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    284 GTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCAGTGGAAA 343

Qy   1918 gattgatcagttcgctgttgcgaacatgtgatttggaagagtttcagacctgtctggtgc 1977
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    344 GATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGTCTGGTGC 403

Qy   1978 ggcactgcaagcatgcgtttggctgtgcgctggtgcacacctctggctggaaagtggtct 2037
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    404 GGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAAGTGGTCT 463

Qy   2038 attccggg 2045
          ||||||||
Db    464 ATTCCGGG 471


RESULT  30
R87541
LOCUS       R87541         518 bp    mRNA            EST       16-AUG-1995
DEFINITION  ym89b04.r1 Soares adult brain N2b4HB55Y Homo sapiens cDNA clone
            IMAGE:166063 5', mRNA sequence.
ACCESSION   R87541
VERSION     R87541.1  GI:946354
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 518)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 2215
            High quality sequence stops: 331 Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 2215   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 331.
FEATURES             Location/Qualifiers
     source          1. .518
                     /organism="Homo sapiens"
                     /db_xref="GDB:586750"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:166063"
                     /clone_lib="Soares adult brain N2b4HB55Y"
                     /sex="Male"
                     /dev_stage="55-year old"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: brain; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 53. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. The adult brain RNA was
                     provided by Dr. Donald H. Gilden. Tissue was acquired
                     17-18 hours after death which occurred in consequence of a
                     ruptured aortic aneurysm. RNA was prepared from a pool of
                     tissues representing the following areas of the brain:
                     frontal, parietal, temporal and occipital cortex from the
                     left and right hemispheres, subcortical white matter,
                     basal ganglia, thalamus, cerebellum, midbrain, pons and
                     medulla."
BASE COUNT      114 a    136 c    137 g    125 t      6 others
ORIGIN


  Query Match             11.9%;  Score 352;  DB 144;  Length 518;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-175;
  Matches  352;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    940 gtactcctccagatcctggtgctgcttttgtggtggtagaatgtccagatgaaagcttca 999
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2 GTACTCCTCCAGATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCA 61

Qy   1000 ttcaacccatctgtgagaatgccacctttcagaggtaccaaggaaaggcagatgccccccg 1059
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     62 TTCAACCCATCTGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCG 121

Qy   1060 tggccttggtggttcacatggccccagcatctgtgcttgtggacagcaggtaccagcagt 1119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 TGGCCTTGGTGGTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGT 181

Qy   1120 ggatggagaggtttgggcctgacacccagcacttggtcctgaatgagaactgtgcctcag 1179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182 GGATGGAGAGGTTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAG 241

Qy   1180 ttcacaaccttcgcagccacaagattcaaacccagctcaacctcatccacccggacatct 1239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    242 TTCACAACCTTCGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCT 301

Qy   1240 tccccctgctcaccagtttccgctgtaagaaggagggccccaccctcagtgt 1291
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    302 TCCCCCTGCTCACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGT 353


RESULT  31
W27286/c
LOCUS       W27286         855 bp    mRNA            EST       08-MAY-1996
DEFINITION  28h1 Human retina cDNA randomly primed sublibrary Homo sapiens cDNA
            , mRNA sequence.
ACCESSION   W27286
VERSION     W27286.1  GI:1307081
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 855)
  AUTHORS   Macke, J., Smallwood, P. and Nathans, J.
  TITLE     Adult Human Retina cDNA
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Dr. Jeremy Nathans
            Dr. Jeremy Nathans, Dept. of Molecular Biology and Genetics
            Johns Hopkins School of Medicine
            725 North Wolfe Street, Baltimore, MD 21205
```

```
            Tel: 410 955 4678
            Fax: 410 614 0827
            Email: jeremy_nathans@qmail.bs.jhu.edu
            Clones from this library are NOT available.
            PCR PRimers
            FORWARD: CTTTTGAGCAAGTTCAGCCTGGTTAAGT
            BACKWARD: GAGGTGGCTTATGAGTATTTCTTCCAGGGTAA
            Seq primer: GGGTAAAAAGCAAAAGAATT.
FEATURES              Location/Qualifiers
     source           1. .855
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone_lib="Human retina cDNA randomly primed sublibrary"
                      /sex="mixed (males and females)"
                      /tissue_type="retina"
                      /dev_stage="adult"
                      /lab_host="E. coli strain K802"
                      /note="Organ: eye; Vector: lambda gt10; Site_1: EcoRI;
                      Site_2: EcoRI; The library used for sequencing was a
                      sublibrary derived from a human retina cDNA library.
                      Inserts from retina cDNA library DNA were isolated,
                      randomly primed, PCR amplified, size-selected, and cloned
                      into lambda gt10.  Individual plaques were arrayed and
                      used as templates for PCR amplification, and these PCR
                      products were used for sequencing."
BASE COUNT      130 a    142 c    162 g    150 t    271 others
ORIGIN


  Query Match             11.9%;  Score 351;  DB 146;  Length 855;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-174;
  Matches  351;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     995 cttcattcaacccatctgtgagaatgccacctttcagaggtaccaaggaaaggcagatgc 1054
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     429 CTTCATTCAACCCATCTGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGC 370

Qy    1055 ccccgtggccttggtggttcacatggccccagcatctgtgcttgtggacagcaggtacca 1114
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     369 CCCCGTGGCCTTGGTGGTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCA 310

Qy    1115 gcagtggatggagaggtttgggcctgacacccagcacttggtcctgaatgagaactgtgc 1174
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     309 GCAGTGGATGGAGAGGTTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGC 250

Qy    1175 ctcagttcacaaccttcgcagccacaagattcaaacccagctcaacctcatccacccgga 1234
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     249 CTCAGTTCACAACCTTCGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGA 190

Qy    1235 catcttccccctgctcaccagtttccgctgtaagaaggagggccccacctcagtgtgcc 1294
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     189 CATCTTCCCCCTGCTCACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCC 130

Qy    1295 catggttcagggtgaatgcctcctcaagtaccagctccgtcccaggaggga 1345
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     129 CATGGTTCAGGGTGAATGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGA 79


RESULT  32
BE892893
LOCUS       BE892893      790 bp    mRNA            EST       29-SEP-2000
DEFINITION  601435738F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3920792 5',
            mRNA sequence.
ACCESSION   BE892893
VERSION     BE892893.1  GI:10353525
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 790)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM739  row: l  column: 09
            High quality sequence stop: 662.
FEATURES              Location/Qualifiers
     source           1. .790
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:3920792"
                      /clone_lib="NIH_MGC_72"
                      /tissue_type="melanotic melanoma"
                      /lab_host="DH10B (phage-resistant)"
                      /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                      Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                      Average insert size 2 kb.   Library constructed by Life
                      Technologies."
BASE COUNT      185 a    221 c    220 g    164 t
ORIGIN


  Query Match             11.7%;  Score 347;  DB 137;  Length 790;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-172;
  Matches  347;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     953 tcctggtgctgcttttgtggtggtagaatgtccagatgaaagcttcattcaacccatctg 1012
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATCTG 60

Qy    1013 tgagaatgccacctttcagaggtaccaaggaaaggcagatgccccctgtggccttggtggt 1072
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTGGT 120

Qy    1073 tcacatggccccagcatctgtgcttgtggacagcaggtaccagcagtggatggagaggtt 1132
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGGTT 180

Qy    1133 tgggcctgacacccagcacttggtcctgaatgagaactgtgcctcagttcacaaccttcg 1192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTTCG 240

Qy    1193 cagccacaagattcaaacccagctcaacctcatccacccggacatcttcccccctgctcac 1252
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTCAC 300

Qy    1253 cagtttccgctgtaagaaggagggccccacctcagtgtgcccatgg 1299
           |||||||||||||||||||||||||||||||||||||||||||||
Db     301 CAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGG 347


RESULT  33
AA811170/c
LOCUS       AA811170      396 bp    mRNA            EST       19-FEB-1998
DEFINITION  ob42c03.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1334020 3',
            mRNA sequence.
ACCESSION   AA811170
VERSION     AA811170.1  GI:2880781
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 396)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
```

```
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 885   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 377.
FEATURES             Location/Qualifiers
     source          1. .396
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1334020"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER).  cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTT-3'
                     ]. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT       81 a    102 c     97 g    116 t
ORIGIN


  Query Match             11.6%;  Score 343;  DB 12;  Length 396;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-170;
  Matches  343;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2616 ctgaagagcacggtcccccaggaggcagctcaggataggtggtatggagctgtgccgagg 2675
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     355 CTGAAGAGCACGGTCCCCCAGGAGGCAGCTCAGGATAGGTGGTATGGAGCTGTGCCGAGG 296

Qy    2676 cttgggctcccacataagcactagtctatagatgcctcttaggactggtgcctggcacag 2735
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     295 CTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTTAGGACTGGTGCCTGGCACAG 236

Qy    2736 ccgcgggccaggaggctgccacacggaagcaagcagatgaactaatttcatttcaaggca 2795
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     235 CCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGAACTAATTTCATTTCAAGGCA 176

Qy    2796 gtttttaaagaagtcttggaaacagacggcggcacctttcctctaatccagcaaagtgat 2855
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     175 GTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTCCTCTAATCCAGCAAAGTGAT 116

Qy    2856 tccctgcacaccagagacaagcagagtaacaggatcagtgggtctaagtgtccgagactt 2915
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     115 TCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTGGGTCTAAGTGTCCGAGACTT 56

Qy    2916 aacgaaaatagtatttcagctgcaataaagattgagtttgcaa 2958
           |||||||||||||||||||||||||||||||||||||||||||
Db      55 AACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTGCAA 13


RESULT  34
AA634909
LOCUS       AA634909      446 bp    mRNA            EST       21-OCT-1997
DEFINITION  ab27h02.r1 Stratagene lung (#937210) Homo sapiens cDNA clone
            IMAGE:842067 5' similar to SW:YK59_YEAST P36159 HYPOTHETICAL 96.8
            KD PROTEIN IN SIS2-MTD1 INTERGENIC REGION. ;, mRNA sequence.
ACCESSION   AA634909
VERSION     AA634909.1  GI:2558123
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 446)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK.
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 430.
FEATURES             Location/Qualifiers
     source          1. .446
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:842067"
                     /clone_lib="Stratagene lung (#937210)"
                     /sex="male"
                     /dev_stage="72 years"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                     dT. normal lung.  Average insert size: 1.0 kb; Uni-ZAP XR
                     Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3'
                     adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT      105 a    112 c    131 g     97 t      1 others
ORIGIN


  Query Match             11.6%;  Score 342;  DB 10;  Length 446;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-169;
  Matches  342;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1944 tgtgatttggaagagtttcagacctgtctggtgcggcactgcaagcatgcgtttggctgt 2003
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      59 TGTGATTTGGAAGAGTTTCAGACCTGTCTGGTGCGGCACTGCAAGCATGCGTTTGGCTGT 118

Qy    2004 gcgctggtgcacacctctggctggaaagtggtctattccggggacaccatgccctgcgag 2063
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     119 GCGCTGGTGCACACCTCTGGCTGGAAAGTGGTCTATTCCGGGGACACCATGCCCTGCGAG 178

Qy    2064 gctctggtccggatggggaaagatgccaccctcctgatacatgaagccaccctggaagat 2123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     179 GCTCTGGTCCGGATGGGGAAAGATGCCACCCTCCTGATACATGAAGCCACCCTGGAAGAT 238

Qy    2124 ggtttggaagaggaagcagtggaaaagacacacagcacaacgtcccaagccatcagcgtg 2183
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     239 GGTTTGGAAGAGGAAGCAGTGGAAAAGACACACAGCACAACGTCCCAAGCCATCAGCGTG 298

Qy    2184 gggatgcggatgaacgcggagttcattatgctgaaccacttcagccagcgctatgccaag 2243
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     299 GGGATGCGGATGAACGCGGAGTTCATTATGCTGAACCACTTCAGCCAGCGCTATGCCAAG 358

Qy    2244 gtccccctcttcagccccaacttcagcgagaaagtgggagtt 2285
           |||||||||||||||||||||||||||||||||||||||||||
Db     359 GTCCCCCTCTTCAGCCCCAACTTCAGCGAGAAAGTGGGAGTT 400
```

```
RESULT  35
AA679618
LOCUS       AA679618      452 bp    mRNA            EST       02-DEC-1997
DEFINITION  ag72c12.s1 Gessler Wilms tumor Homo sapiens cDNA clone
            IMAGE:1128502 3' similar to SW:YK59_YEAST P36159 HYPOTHETICAL 96.8
            KD PROTEIN IN SIS2-MTD1 INTERGENIC REGION. ;, mRNA sequence.
ACCESSION   AA679618
VERSION     AA679618.1  GI:2660140
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 452)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: similarity on wrong strand
            Possible reversed clone: polyT not found
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .452
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1128502"
                     /clone_lib="Gessler Wilms tumor"
                     /sex="pooled (6)"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; RNA
                     was prepared from a pool of 6 anonymous Wilms' tumor RNAs.
                     RNA was prepared by acid-phenol, followed by one round of
                     oligo dT selection.  cDNA library preparation was with
                     the BRL/Life Tech. Superscript Plasmid system.  An
                     oligo-dT NotI primer for first strand synthesis generated
                     gcggccgccc(t)n at the 3' end of the clones.  A 5' SalI
                     adaptor was used  with sequence 5'-gtcgacccacgcgtccg-3'.
                     Resulting cDNAs were size selected (average size 2 kb),
                     NotI digested, and ligated into NotI/SalI-cut pSPORT1.
                     Library was constructed by Dr. Manfred Gessler."
BASE COUNT      111 a    120 c    126 g     95 t
ORIGIN
```

```
  Query Match            11.6%;  Score 342;  DB 10;  Length 452;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-169;
  Matches  342;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    251 ccgggactcgggcgccgcgctctacgtcttctccgagttcaaccggtatctcttcaactg 310
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    111 CCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTCAACCGGTATCTCTTCAACTG 170

Qy    311 tggagaaggcgttcagagactcatgcaggagcacaagttaaaggttgctcgcctggacaa 370
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    171 TGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTAAAGGTTGCTCGCCTGGACAA 230

Qy    371 catattcctgacacgaatgcactggtctaatgttgggggcttaagtggaatgattcttac 430
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    231 CATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGCTTAAGTGGAATGATTCTTAC 290

Qy    431 tttaaaggaaaccgggcttccaaagtgtgtactttctggacctccacaactggaaaaata 490
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    291 TTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATA 350

Qy    491 cctcgaagcaatcaaaatattttctggtccattgaaaggaatagaactggctgtgcggcc 550
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    351 CCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGAATAGAACTGGCTGTGCGGCC 410

Qy    551 ccactctgccccagaatacgaggatgaaaccatgacagttta 592
          |||||||||||||||||||||||||||||||||||||||||||
Db    411 CCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTTTA 452
```

```
RESULT  36
BE795434
LOCUS       BE795434      698 bp    mRNA            EST       20-SEP-2000
DEFINITION  601592991F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3946774 5',
            mRNA sequence.
ACCESSION   BE795434
VERSION     BE795434.1  GI:10216632
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 698)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM807  row: f  column: 23
            High quality sequence stop: 389.
FEATURES             Location/Qualifiers
     source          1. .698
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3946774"
                     /clone_lib="NIH_MGC_7"
                     /tissue_type="small cell carcinoma"
                     /cell_line="MGC3"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      201 a    212 c    175 g    110 t
ORIGIN
```

```
  Query Match            11.5%;  Score 341;  DB 135;  Length 698;
  Best Local Similarity  99.7%;  Pred. No. 5e-169;
  Matches  391;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   1529 gatgaagattcgaaatgtcagtgccacacttgtcaacataagccccgacacgtctctgct 1588
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GATGAAGATTCGAAATGTCAGTGCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCT 60

Qy   1589 actggactgtggtgagggcacatttgggcagctgtgccgtcattacggagaccaggtgga 1648
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     61 ACTGGACTGTGGTGAGGGCACGTTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGA 120

Qy   1649 cagggtcctgggcaccctggctgctgtgtttgtgtcccacctgcacgcagatcaccacac 1708
```

```
Db    121 CAGGGTCCTGGGCACCCTGGCTGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACAC 180

Qy   1709 gggcttgccaagtatcttgctgcagagagaacgcgccttggcatctttgggaaagccgct 1768
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGGCTTGCCAAGTATCTTGCTGCAGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCT 240

Qy   1769 tcaccctttgctggtggttgcccccaaccagctcaaagcctggctccagcagtaccacaa 1828
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCACCCTTTGCTGGTGGTTGCCCCCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAA 300

Qy   1829 ccagtgccaggaggtcctgcaccacatcagtatgattcctgccaaatgccttcaggaagg 1888
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CCAGTGCCAGGAGGTCCTGCACCACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGG 360

Qy   1889 ggctgagatctccagtcctgcagtggaaagat 1920
          ||||||||||||||||||||||||||||||
Db    361 GGCTGAGATCTCCAGTCCTGCAGTGGAAAGAT 392


RESULT  37
AW407520/c
LOCUS       AW407520      345 bp    mRNA            EST       16-FEB-2000
DEFINITION  UI-HF-BM0-adl-b-07-0-UI.r1 NIH_MGC_38 Homo sapiens cDNA clone
            IMAGE:3061957 5', mRNA sequence.
ACCESSION   AW407520
VERSION     AW407520.1  GI:6926577
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 345)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Eco RI site shown at the beginning of the sequence.
             Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: M.B. Soares Lab
             cDNA Library Arrayed by: M.B. Soares Lab
             DNA Sequencing by: M.B. Soares Lab
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
     source          1. .345
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3061957"
                     /clone_lib="NIH_MGC_38"
                     /tissue_type="lymph"
                     /cell_type="germinal center B cells"
                     /cell_line="MGC85"
                     /lab_host="DH10B (LTI)"
                     /note="Vector: pT7T3-Pac; Site_1: NotI; Site_2: Eco RI;
                     Constructed from size fractionated cytoplasmic mRNA
                     (2.5-3.5kb). Directionally cloned. Cells provided by Louis
                     M. Staudt, Ph.D. Library preparation by Maria de Fatima
                     Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
BASE COUNT       73 a     94 c     82 g     96 t
ORIGIN


  Query Match             11.4%;  Score 338;  DB 90;  Length 345;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-167;
  Matches  338;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2619 aagagcacggtcccccaggaggcagctcaggataggtggtatggagctgtgccgaggctt 2678
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    345 AAGAGCACGGTCCCCCAGGAGGCAGCTCAGGATAGGTGGTATGGAGCTGTGCCGAGGCTT 286

Qy   2679 gggctcccacataagcactagtctatagatgcctcttaggactggtgcctggcacagccg 2738
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    285 GGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTTAGGACTGGTGCCTGGCACAGCCG 226

Qy   2739 cgggccaggaggctgccacacggaagcaagcagatgaactaatttcatttcaaggcagtt 2798
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    225 CGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGAACTAATTTCATTTCAAGGCAGTT 166

Qy   2799 tttaaagaagtcttggaaacagacggcggcacctttcctctaatccagcaaagtgattcc 2858
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    165 TTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTCCTCTAATCCAGCAAAGTGATTCC 106

Qy   2859 ctgcacaccagagacaagcagagtaacaggatcagtgggtctaagtgtccgagacttaac 2918
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    105 CTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAAC 46

Qy   2919 gaaaatagtatttcagctgcaataaagattgagtttgc 2956
          |||||||||||||||||||||||||||||||||||
Db     45 GAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTGC 8


RESULT  38
BE742908
LOCUS       BE742908      677 bp    mRNA            EST       15-SEP-2000
DEFINITION  601574609F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3835658 5',
            mRNA sequence.
ACCESSION   BE742908
VERSION     BE742908.1  GI:10156900
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 677)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM518  row: a  column: 03
            High quality sequence stop: 672.
FEATURES             Location/Qualifiers
     source          1. .677
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3835658"
                     /clone_lib="NIH_MGC_9"
                     /tissue_type="adenocarcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      160 a    173 c    217 g    127 t
ORIGIN


  Query Match             11.3%;  Score 334;  DB 135;  Length 677;
  Best Local Similarity   99.8%;  Pred. No. 2.6e-165;
  Matches  454;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;
```

```
Qy   1857 agtatgattcctgccaaatgccttcaggaaggggctgagatctccagtcctgcagtggaa 1916
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCAGTGGAA 60

Qy   1917 agattgatcagttcgctgttgcgaacatgtgatttggaagagtttcagacctgtctggtg 1976
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGTCTGGTG 120

Qy   1977 cggcactgcaagcatgcgtttggctgtgcgctggtgcacacctctggctggaaagtggtc 2036
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCTGGAAAGTGGTC 180

Qy   2037 tattccggggacaccatgccctgcgaggctctggtccggatggggaaagatgccaccctc 2096
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAGATGCCACCCTC 240

Qy   2097 ctgatacatgaagccaccctggaagatggtttggaagaggaagcagtggaaaagacacac 2156
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    241 CTGATACATGAAGCCACCCTGGAAGATGG-TTGGAAGAGGAAGCAGTGGAAAAGACACAC 299

Qy   2157 agcacaacgtcccaagccatcagcgtggggatgcggatgaacgcggagttcattatgctg 2216
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 AGCACAACGTCCCAAGCCATCAGCGTGGGGATGCGGATGAACGCGGAGTTCATTATGCTG 359

Qy   2217 aaccacttcagccagcgctatgccaaggtccccctcttcagccccaacttcagcgagaaa 2276
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 AACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTCAGCGAGAAA 419

Qy   2277 gtgggagttgcctttgaccacatgaaggtctgctt 2311
          |||||||||||||||||||||||||||||||||||
Db    420 GTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTT 454
```

RESULT 39
BE298273
LOCUS BE298273 416 bp mRNA EST 20-JUL-2000
DEFINITION 601118143F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:3027851 5', mRNA sequence.
ACCESSION BE298273
VERSION BE298273.1 GI:9182013
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 416)
AUTHORS NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM85 row: f column: 12
High quality sequence stop: 416.

```
FEATURES             Location/Qualifiers
     source          1. .416
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3027851"
                     /clone_lib="NIH_MGC_17"
                     /tissue_type="rhabdomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
                     Site_2: XhoI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT       81 a    124 c    132 g     79 t
ORIGIN


  Query Match           11.2%;  Score 330;  DB 106;  Length 416;
  Best Local Similarity  99.7%;  Pred. No. 3.3e-163;
  Matches  380;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     57 gcgctttgctcgctgctgcggtccgcggccggacgcaccatgtcgcagggacgcaccata 116
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGCACCATA 60

Qy    117 tcgcaggcacccgcccgccgcgagcggccgcgcaaggacccgctgcggcacctgcgcacg 176
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTGCGCACG 120

Qy    177 cgagagaagcgcggaccgtcggggtgctccggcggcccaaacaccgtgtacctgcaggtg 236
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTGCAGGTG 180

Qy    237 gtggcagcgggtagccgggactcgggcgccgcgctctacgtcttctccgagttcaaccgg 296
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTCAACCGG 240

Qy    297 tatctcttcaactgtggagaaggcgttcagagactcatgcaggagcacaagttaaaggtt 356
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTAAAGGTT 300

Qy    357 gctcgcctggacaacatattcctgacacgaatgcactggtctaatgttgggggcttaagt 416
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGCTTAAGT 360

Qy    417 ggaatgattcttactttaaag 437
          |||||||||||||||||||||
Db    361 GGAATGATTCTTACTTTAAAG 381
```

RESULT 40
BE743831
LOCUS BE743831 940 bp mRNA EST 15-SEP-2000
DEFINITION 601577742F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3926362 5', mRNA sequence.
ACCESSION BE743831
VERSION BE743831.1 GI:10157823
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 940)
AUTHORS NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM754 row: d column: 11
High quality sequence stop: 631.

```
FEATURES             Location/Qualifiers
     source          1. .940
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3926362"
```

/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
ASE COUNT 250 a 242 c 281 g 167 t
RIGIN

Query Match 11.1%; Score 329; DB 135; Length 940;
Best Local Similarity 100.0%; Pred. No. 1.2e-162;
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 755 ggactcttccctggtcgtagctttcatctgtaagcttcacttaaagagaggaaacttctt 814
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 33 GGACTCTTCCCTGGTCGTAGCTTTCATCTGTAAGCTTCACTTAAAGAGAGGAAACTTCTT 92

Qy 815 ggtgctcaaagcaaaggagatgggcctcccagttgggacagctgccatcgctcccatcat 874
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 93 GGTGCTCAAAGCAAAGGAGATGGGCCTCCCAGTTGGGACAGCTGCCATCGCTCCCATCAT 152

Qy 875 tgctgctgtcaaggacgggaaaagcatcactcatgaaggaagagagattttggctgaaga 934
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 153 TGCTGCTGTCAAGGACGGGAAAAGCATCACTCATGAAGGAAGAGAGATTTTGGCTGAAGA 212

Qy 935 gctgtgtactcctccagatcctggtgctgcttttgtggtggtagaatgtccagatgaaag 994
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 213 GCTGTGTACTCCTCCAGATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAG 272

Qy 995 cttcattcaacccatctgtgagaatgccacctttcagaggtaccaaggaaaggcagatgc 1054
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 273 CTTCATTCAACCCATCTGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGC 332

Qy 1055 ccccgtggccttggtggttcacatggccc 1083
|||||||||||||||||||||||||||||
Db 333 CCCCGTGGCCTTGGTGGTTCACATGGCCC 361

RESULT 41
AI937465/c
LOCUS AI937465 531 bp mRNA EST 08-MAR-2000
DEFINITION wp77e01.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2467800 3', mRNA sequence.
ACCESSION AI937465
VERSION AI937465.1 GI:5676335
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 531)
AUTHORS NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGAP), Tumor Gene Index
JOURNAL Unpublished (1998)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1291 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 467.
FEATURES Location/Qualifiers
source 1. .531
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2467800"
/clone_lib="NCI_CGAP_Brn25"
/tissue_type="anaplastic oligodendroglioma"
/lab_host="DH10B"
/note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCATAGGTTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 100 a 154 c 134 g 143 t
ORIGIN

Query Match 11.0%; Score 325; DB 27; Length 531;
Best Local Similarity 99.2%; Pred. No. 1.5e-160;
Matches 525; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2430 ctcctgtccagggagctggcaggcggcctggaggatggggagcctcagcagaagcgggcc 2489
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 531 CTCCTGTCCAGGGAGCTGGCAGGCGGCCTGGAGGATGGGGAGCCTCAGCAGAAGCGGGCC 472

Qy 2490 cacacagaggagccacaggccaagaaggtcagagcccagtgaagatctgggagaccctga 2549
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 471 CACACAGAGGAGCCACAGGCCAAGAAGGTCAGAGCCCAGTGAAGATCTGGGAGACCCTGA 412

Qy 2550 actcagaaggctgtgtgtcttctgccccacgcacgcacccgtatctgccctccttgctgg 2609
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 411 ACTCAGAAGGCTGTGTGTCTTCTGCCCCACGCACGCACCCGTATCTGCCCTCCTTGCTGG 352

Qy 2610 tagaagctgaagagcacggtcccccaggaggcagctcaggataggtggtatggagctgtg 2669
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 351 TAGAAGCTGAAGAGCACGGTCCCCCAGGAGGCAGCTCAGGATAGGTGGTATGGAGCTGTG 292

Qy 2670 ccgaggcttgggctcccacataagcactagtctatagatgcctcttaggactggtgcctg 2729
|||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 291 CCGAGGCTTGGGGTCCCACATAAGCACTAGTCTATAGATGCCTCTTAGGACTGGTGCCTG 232

Qy 2730 gcacagccgcgggccaggaggctgccacacggaagcaagcagatgaactaatttcatttc 2789
||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 231 GCACAGCTGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGAACTAATTTCATTTC 172

Qy 2790 aaggcagtttttaaagaagtcttggaaacagacggcggcaccttttcctctaatccagcaa 2849
|||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db 171 AAGGCAGTTTTTAAAGAAGTCATGGAAACAGACGGCGGCACCTTTCCTCTAATCCAGCAA 112

Qy 2850 agtgattccctgcacaccagagacaagcagagtaacaggatcagtgggtctaagtgtccg 2909
| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 111 AATGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTGGGTCTAAGTGTCCG 52

Qy 2910 agacttaacgaaaatagtatttcagctgcaataaagattgagtttgcaa 2958
|||||||||||||||||||||||||||||||||||||||||||||||||||
Db 51 AGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTGCAA 3

RESULT 42
AA632118
LOCUS AA632118 397 bp mRNA EST 28-OCT-1997
DEFINITION np66h03.s1 NCI_CGAP_Br2 Homo sapiens cDNA clone IMAGE:1131317 3' similar to SW:YK59_YEAST P36159 HYPOTHETICAL 96.8 KD PROTEIN IN SIS2-MTD1 INTERGENIC REGION. ;, mRNA sequence.
ACCESSION AA632118

```
VERSION     AA632118.1  GI:2555532
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 397)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1436   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 383.
FEATURES              Location/Qualifiers
     source           1. .397
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1131317"
                      /clone_lib="NCI_CGAP_Br2"
                      /sex="female, pooled"
                      /tissue_type="breast"
                      /lab_host="DH10B"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; 1st strand cDNA was prepared from pooled bulk
                      breast tumor tissue, and was then primed with a Not I -
                      oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                      RI adaptors (Pharmacia), digested with Not I and cloned
                      into the Not I and Eco RI sites of the modified pT7T3
                      vector.  This library is the normalized version of
                      NCI_CGAP_Br1.1. Library was constructed by Bento Soares
                      and M. Fatima Bonaldo. "
BASE COUNT       98 a    100 c    111 g     88 t
ORIGIN
```

```
  Query Match             10.9%;  Score 323;  DB 10;  Length 397;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-159;
  Matches  323;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    251 ccgggactcgggcgccgcgctctacgtcttctccgagttcaaccggtatctcttcaactg 310
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     71 CCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTCAACCGGTATCTCTTCAACTG 130

Qy    311 tggagaaggcgttcagagactcatgcaggagcacaagttaaaggttgctcgcctggacaa 370
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    131 TGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTAAAGGTTGCTCGCCTGGACAA 190

Qy    371 catattcctgacacgaatgcactggtctaatgttgggggcttaagtggaatgattcttac 430
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    191 CATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGCTTAAGTGGAATGATTCTTAC 250

Qy    431 tttaaaggaaaccgggcttccaaagtgtgtactttctggacctccacaactggaaaaata 490
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    251 TTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATA 310

Qy    491 cctcgaagcaatcaaaatattttctggtccattgaaaggaatagaactggctgtgcggcc 550
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    311 CCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGAATAGAACTGGCTGTGCGGCC 370

Qy    551 ccactctgccccagaatacgagg 573
          |||||||||||||||||||||||
Db    371 CCACTCTGCCCCAGAATACGAGG 393
```

```
RESULT  43
R55841
LOCUS       R55841         482 bp    mRNA            EST       23-MAY-1995
DEFINITION  yg89d01.r1 Soares infant brain 1NIB Homo sapiens cDNA clone
            IMAGE:40931 5' similar to SP:YK59_YEAST P36159 HYPOTHETICAL 96.8 KD
            PROTEIN IN SIS2-MTD1 INTERGENIC ;, mRNA sequence.
ACCESSION   R55841
VERSION     R55841.1  GI:825947
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 482)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1777
            High quality sequence stops: 387 Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1777   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 387.
FEATURES              Location/Qualifiers
     source           1. .482
                      /organism="Homo sapiens"
                      /db_xref="GDB:413472"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:40931"
                      /clone_lib="Soares infant brain 1NIB"
                      /sex="female"
                      /dev_stage="73 days post natal"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Organ: whole brain; Vector: Lafmid BA; Site_1: Not
                      I; Site_2: Hind III; 1st strand cDNA was primed with a Not
                      I - oligo(dT) primer [5'
                      AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'];
                      double-stranded cDNA was ligated to Hind III adaptors
                      (Pharmacia), digested with Not I and directionally cloned
                      into the Not I and Hind III sites of the Lafmid BA vector.
                      Library went through one round of normalization. Library
                      constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      107 a    139 c    131 g    102 t      3 others
ORIGIN
```

```
  Query Match             10.9%;  Score 323;  DB 144;  Length 482;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-159;
  Matches  323;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1430 gcaggagtacaggaggagtgcgcaggacggcccagccccagcagagaaaagaagtcagta 1489
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     43 GCAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTA 102

Qy   1490 cccagaaatcatcttccttggaacagggtctgccatcccgatgaagattcgaaatgtcag 1549
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    103 CCCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAG 162

Qy   1550 tgccacacttgtcaacataagccccgacacgtctctgctactggactgtggtgagggcac 1609
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    163 TGCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCAC 222

Qy   1610 atttgggcagctgtgccgtcattacggagaccaggtggacagggtcctgggcaccctggc 1669
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    223 ATTTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGC 282

Qy   1670 tgctgtgtttgtgtcccacctgcacgcagatcaccacacgggcttgccaagtatcttgct 1729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    283 TGCTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCT 342

Qy   1730 gcagagagaacgcgccttggcat 1752
          |||||||||||||||||||||||
Db    343 GCAGAGAGAACGCGCCTTGGCAT 365


RESULT  44
AA233087/c
LOCUS       AA233087      366 bp    mRNA            EST       28-FEB-1997
DEFINITION  zr68g08.r1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:668606
            5', mRNA sequence.
ACCESSION   AA233087
VERSION     AA233087.1  GI:1856275
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 366)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 346.
FEATURES             Location/Qualifiers
     source          1. .366
                     /organism="Homo sapiens"
                     /db_xref="GDB:5562573"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:668606"
                     /clone_lib="Soares_NhHMPu_S1"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
BASE COUNT       83 a     90 c    110 g     83 t
ORIGIN


  Query Match           10.6%;  Score 315;  DB 4;  Length 366;
  Best Local Similarity  99.7%;  Pred. No. 2.9e-155;
  Matches  365;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    931 aagagctgtgtactcctccagatcctggtgctgcttttgtggtggtagaatgtccagatg 990
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    366 AAGAGCTGTGTACTCCTCCAGATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATG 307

Qy    991 aaagcttcattcaacccatctgtgagaatgccacctttcagaggtaccaaggaaaggcag 1050
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    306 AAAGCTTCATTCAACCCATCTGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAATGCAG 247

Qy   1051 atgcccccgtggccttggtggttcacatggccccagcatctgtgcttgtggacagcaggt 1110
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    246 ATGCCCCCGTGGCCTTGGTGGTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGT 187

Qy   1111 accagcagtggatggagaggtttgggcctgacacccagcacttggtcctgaatgagaaact 1170
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    186 ACCAGCAGTGGATGGAGAGGTTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACT 127

Qy   1171 gtgcctcagttcacaaccttcgcagccacaagattcaaacccagctcaacctcatccacc 1230
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    126 GTGCCTCAGTTCACAACCTTCGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACC 67

Qy   1231 cggacatcttccccctgctcaccagtttccgctgtaagaaggagggccccacccctcagtg 1290
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     66 CGGACATCTTCCCCCTGCTCACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTG 7

Qy   1291 tgccca 1296
          ||||||
Db      6 TGCCCA 1


RESULT  45
BE902696
LOCUS       BE902696      735 bp    mRNA            EST       29-SEP-2000
DEFINITION  601677393F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3959926 5',
            mRNA sequence.
ACCESSION   BE902696
VERSION     BE902696.1  GI:10393148
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 735)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM841  row: j  column: 23
            High quality sequence stop: 732.
FEATURES             Location/Qualifiers
     source          1. .735
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3959926"
                     /clone_lib="NIH_MGC_21"
                     /tissue_type="choriocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb.  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      168 a    206 c    203 g    157 t      1 others
```

```
ORIGIN

  Query Match            10.4%;  Score 307;  DB 137;  Length 735;
  Best Local Similarity  100.0%;  Pred. No. 5.2e-151;
  Matches  307;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1372 gcaatcctgaggaattcatagttgaggcgctgcagcttcccaacttccagcagagcgtgc 1431
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    134 GCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCCAGCAGAGCGTGC 193

Qy   1432 aggagtacaggaggagtgcgcaggacggcccagccccagcagagaaaagaagtcagtacc 1491
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    194 AGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTACC 253

Qy   1492 cagaaatcatcttccttggaacagggtctgccatcccgatgaagattcgaaatgtcagtg 1551
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    254 CAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAGTG 313

Qy   1552 ccacacttgtcaacataagccccgacacgtctctgctactggactgtggtgagggcacat 1611
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    314 CCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACAT 373

Qy   1612 ttgggcagctgtgccgtcattacggagaccaggtggacagggtcctgggcaccctggctg 1671
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    374 TTGGGCAGCTGTGCCGTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCTG 433

Qy   1672 ctgtgtt 1678
          |||||||
Db    434 CTGTGTT 440


RESULT  46
AA676661/c
LOCUS       AA676661      491 bp    mRNA            EST       19-DEC-1997
DEFINITION  zj67h01.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA
            clone IMAGE:459985 3', mRNA sequence.
ACCESSION   AA676661
VERSION     AA676661.1  GI:2657183
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 491)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES             Location/Qualifiers
     source          1..491
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:459985"
                     /clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     This is a subtracted version of the original Soares fetal
                     liver spleen 1NFLS library.  1st strand cDNA was primed
                     with a Pac I - oligo(dT) primer [5'
                     AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       94 a    139 c    125 g    133 t
ORIGIN


  Query Match            10.2%;  Score 302;  DB 10;  Length 491;
  Best Local Similarity  99.3%;  Pred. No. 2.3e-148;
  Matches  452;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy   2502 ccacaggccaagaaggtcagagcccagtgaagatctgggagaccctgaactcagaaggct 2561
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    455 CCACAGGCCAAGAAGGTCAGAGCCCAGTGAAGATCTGGGAGACCCTGAACTCAGAAGGCT 396

Qy   2562 gtgtgtcttctgccccacgcacgcacccgtatctgccctccttgctggtagaagctgaag 2621
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    395 GTGTGTCTTCTGCCCCACGCACGCACCCGTATCTGCCCTCCTTGCTGGTAGAAGCTGAAG 336

Qy   2622 agcacggtcccccaggaggcagctcaggataggtggtatggagctgtgccgaggcttggg 2681
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    335 AGCACGGTCCCCCAGGAGGCAGCTCAGGATAGGTGGTATGGAGCTGTGCCGAGGCTTGGG 276

Qy   2682 ctcccacataagcactagtctatagatgcctcttaggactggtgcctggcacagccgcgg 2741
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db    275 GTCCCACATAAGCACTAGTCTATAGATGCCTCTTAGGACTGGTGCCTGGCACAGCTGCGG 216

Qy   2742 gccaggaggctgccacacggaagcaagcagatgaactaatttcatttcaaggcagttttt 2801
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    215 GCCAGGAGGCTGCCACACGGAAGCAAGCAGATGAACTAATTTCATTTCAAGGCAGTTTTT 156

Qy   2802 aaagaagtcttggaaacagacggcggcacctttcctctaatccagcaaagtgattccctg 2861
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    155 AAAGAAGTCATGGAAACAGACGGCGGCACCTTTCCTCTAATCCAGCAAAGTGATTCCCTG 96

Qy   2862 cacaccagagacaagcagagtaacaggatcagtgggtctaagtgtccgagacttaacgaa 2921
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     95 CACACCAGAGACAAGCAGAGTAACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACGAA 36

Qy   2922 aatagtatttcagctgcaataaagattgagtttgc 2956
          |||||||||||||||||||||||||||||||||
Db     35 AATAGTATTTCAGCTGCAATAAAGATTGAGTTTGC 1


RESULT  47
R90875
LOCUS       R90875        480 bp    mRNA            EST       25-AUG-1995
DEFINITION  yn01d02.r1 Soares adult brain N2b4HB55Y Homo sapiens cDNA clone
            IMAGE:167139 5', mRNA sequence.
ACCESSION   R90875
VERSION     R90875.1  GI:958415
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 480)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
```

```
           Insert Size: 4441
           High quality sequence stops: 263 Source: IMAGE Consortium, LLNL
           This clone is available royalty-free through LLNL ; contact the
           IMAGE Consortium (info@image.llnl.gov) for further information.
           Insert Length: 4441   Std Error: 0.00
           Seq primer: M13RP1
           High quality sequence stop: 263.
FEATURES             Location/Qualifiers
     source          1. .480
                     /organism="Homo sapiens"
                     /db_xref="GDB:587834"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:167139"
                     /clone_lib="Soares adult brain N2b4HB55Y"
                     /sex="Male"
                     /dev_stage="55-year old"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: brain; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 53. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. The adult brain RNA was
                     provided by Dr. Donald H. Gilden. Tissue was acquired
                     17-18 hours after death which occurred in consequence of a
                     ruptured aortic aneurysm. RNA was prepared from a pool of
                     tissues representing the following areas of the brain:
                     frontal, parietal, temporal and occipital cortex from the
                     left and right hemispheres, subcortical white matter,
                     basal ganglia, thalamus, cerebellum, midbrain, pons and
                     medulla."
BASE COUNT      105 a    125 c    125 g    117 t      8 others
ORIGIN


  Query Match            10.2%;  Score 301;  DB 144;  Length 480;
  Best Local Similarity  100.0%;  Pred. No. 7.7e-148;
  Matches  301;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    940 gtactcctccagatcctggtgctgcttttgtggtggtagaatgtccagatgaaagcttca 999
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTACTCCTCCAGATCCTGGTGCTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCA 60

Qy   1000 ttcaacccatctgtgagaatgccacctttcagaggtaccaaggaaaggcagatgcccccg 1059
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCAACCCATCTGTGAGAATGCCACCTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCG 120

Qy   1060 tggccttggtggttcacatggccccagcatctgtgcttgtggacagcaggtaccagcagt 1119
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGGCCTTGGTGGTTCACATGGCCCCAGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGT 180

Qy   1120 ggatggagaggtttgggcctgacacccagcacttggtcctgaatgagaactgtgcctcag 1179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGATGGAGAGGTTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAG 240

Qy   1180 ttcacaaccttcgcagccacaagattcaaacccagctcaacctcatccacccggacatct 1239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TTCACAACCTTCGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCT 300

Qy   1240 t 1240
          |
Db    301 T 301


RESULT  48
AW771657
LOCUS       AW771657      511 bp    mRNA            EST       04-MAY-2000
DEFINITION  hn59h06.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3032219 3'
            similar to SW:YK59_YEAST P36159 HYPOTHETICAL 96.8 KD PROTEIN IN
            SIS2-MTD1 INTERGENIC REGION. ;, mRNA sequence.
ACCESSION   AW771657
VERSION     AW771657.1  GI:7703715
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 511)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Possible reversed clone: similarity on wrong strand
            Possible reversed clone: polyT not found
            Seq primer: -40UP from Gibco
            High quality sequence stop: 494.
FEATURES             Location/Qualifiers
     source          1. .511
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3032219"
                     /clone_lib="NCI_CGAP_Kid11"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo.  "
BASE COUNT      132 a    138 c    141 g    100 t
ORIGIN


  Query Match            10.2%;  Score 301;  DB 95;  Length 511;
  Best Local Similarity  99.6%;  Pred. No. 7.7e-148;
  Matches  471;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy    179 agagaagcgcggaccgtcggggtgctccggcggcccaaacaccgtgtacctgcaggtggt 238
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     40 AGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCGTGTACCTGCAGGTGGT 99

Qy    239 ggcagcgggtagccgggactcgggcgccgcgctctacgtcttctccgagttcaaccggta 298
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100 GGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTCAACCGGTA 159

Qy    299 tctcttcaactgtggagaaggcgttcagagactcatgcaggagcacaagttaaaggttgc 358
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    160 TCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTAAAGGTTGC 219

Qy    359 tcgcctggacaacatattcctgacacgaatgcactggtctaatgttgggggcttaagtgg 418
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    220 TCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGCTTAAGTGG 279

Qy    419 aatgattcttactttaaaggaaaccgggcttccaaagtgtgtactttctggacctccaca 478
          ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db    280 AATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTG-CTACTTTCTGGACCTCCACA 338
```

```
Qy    479 actggaaaaatacctcgaagcaatcaaaatattttctggtccattgaaaggaatagaact 538
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    339 ACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGAATAGAACT 398

Qy    539 ggctgtgcggccccactctgccccagaatacgaggatgaaaccatgacagtttaccagat 598
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    399 GGCTGTGCGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTTTACCAGAT 458

Qy    599 ccccatacacagtgaacagaggaggggaaagcaccaaccatggcagagtccag 651
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    459 CCCCATACACAGTGAACAGAGGAGGGGAAAGCACCAACCATGGCAGAGTCCAG 511


RESULT 49
AW994476
LOCUS       AW994476      297 bp    mRNA            EST       05-JUN-2000
DEFINITION  RC3-BN0036-260400-015-e06 BN0036 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW994476
VERSION     AW994476.1  GI:8254699
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 297)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC3-BN0036-260
            400-015-e06&t3=2000-04-26&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 9
            High quality sequence stop: 297.
FEATURES             Location/Qualifiers
     source          1. .297
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="BN0036"
                     /dev_stage="Adult"
                     /note="Organ: breast_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       60 a     93 c     82 g     62 t
ORIGIN


  Query Match             9.8%;  Score 291;  DB 97;  Length 297;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-142;
  Matches  291;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1567 taagccccgacacgtctctgctactggactgtggtgagggcacatttgggcagctgtgcc 1626
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7 TAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCACATTTGGGCAGCTGTGCC 66

Qy    1627 gtcattacggagaccaggtggacagggtcctgggcaccctggctgctgtgtttgtgtccc 1686
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67 GTCATTACGGAGACCAGGTGGACAGGGTCCTGGGCACCCTGGCTGCTGTGTTTGTGTCCC 126

Qy    1687 acctgcacgcagatcaccacacgggcttgccaagtatcttgctgcagagagaacgcgcct 1746
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     127 ACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTGCAGAGAGAACGCGCCT 186

Qy    1747 tggcatctttgggaaagccgcttcacccttgctggtggttgcccccaaccagctcaaag 1806
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     187 TGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCCCCCAACCAGCTCAAAG 246

Qy    1807 cctggctccagcagtaccacaaccagtgccaggaggtcctgcaccacatca 1857
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     247 CCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCACCACATCA 297


RESULT 50
AI174501
LOCUS       AI174501      404 bp    mRNA            EST       07-OCT-1998
DEFINITION  an42e05.s1 Gessler Wilms tumor Homo sapiens cDNA clone
            IMAGE:1701344 3' similar to SW:YK59_YEAST P36159 HYPOTHETICAL 96.8
            KD PROTEIN IN SIS2-MTD1 INTERGENIC REGION. ;, mRNA sequence.
ACCESSION   AI174501
VERSION     AI174501.1  GI:3721354
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 404)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: similarity on wrong strand
            Possible reversed clone: polyT not found
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 391.
FEATURES             Location/Qualifiers
     source          1. .404
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1701344"
                     /clone_lib="Gessler Wilms tumor"
                     /sex="pooled (6)"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; RNA
                     was prepared from a pool of 6 anonymous Wilms' tumor RNAs.
                     RNA was prepared by acid-phenol, followed by one round of
                     oligo dT selection.  cDNA library preparation was with
                     the BRL/Life Tech. Superscript Plasmid system.  An
                     oligo-dT NotI primer for first strand synthesis generated
                     gcggccgccc(t)n at the 3' end of the clones.  A 5' SalI
                     adaptor was used  with sequence 5'-gtcgacccacgcgtccg-3'.
                     Resulting cDNAs were size selected (average size 2 kb),
                     NotI digested, and ligated into NotI/SalI-cut pSPORT1.
                     Library was constructed by Dr. Manfred Gessler."
BASE COUNT       97 a    104 c    115 g     87 t      1 others
ORIGIN
```

```
  Query Match             9.8%;  Score 291;  DB 17;  Length 404;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-142;
  Matches  291;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    255 gactcgggcgccgcgctctacgtcttctccgagttcaaccggtatctcttcaactgtgga 314
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    114 GACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTCAACCGGTATCTCTTCAACTGTGGA 173

Qy    315 gaaggcgttcagagactcatgcaggagcacaagttaaaggttgctcgcctggacaacata 374
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    174 GAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTAAAGGTTGCTCGCCTGGACAACATA 233

Qy    375 ttcctgacacgaatgcactggtctaatgttgggggcttaagtggaatgattcttacttta 434
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    234 TTCCTGACACGAATGCACTGGTCTAATGTTGGGGGCTTAAGTGGAATGATTCTTACTTTA 293

Qy    435 aaggaaaccgggcttccaaagtgtgtactttctggacctccacaactggaaaaatacctc 494
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    294 AAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATACCTC 353

Qy    495 gaagcaatcaaaatattttctggtccattgaaaggaatagaactggctgtg 545
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    354 GAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGAATAGAACTGGCTGTG 404


RESULT  51
BE858252/c
LOCUS       BE858252      494 bp    mRNA            EST       29-SEP-2000
DEFINITION  7g21a09.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:3307096 3',
            mRNA sequence.
ACCESSION   BE858252
VERSION     BE858252.1  GI:10372932
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 494)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
            Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco
            High quality sequence stop: 493.
FEATURES             Location/Qualifiers
     source          1. .494
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3307096"
                     /clone_lib="NCI_CGAP_Brn23"
                     /tissue_type="glioblastoma (pooled)"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATATCTTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT       95 a    140 c    124 g    135 t
ORIGIN


  Query Match             9.8%;  Score 290;  DB 136;  Length 494;
  Best Local Similarity   99.2%;  Pred. No. 5.2e-142;
  Matches  490;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy   2463 gatggggagcctcagcagaagcgggcccacacagaggagccacaggccaagaaggtcaga 2522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    494 GATGGGGAGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGCCAAGAAGGTCAGA 435

Qy   2523 gcccagtgaagatctgggagaccctgaactcagaaggctgtgtgtcttctgccccacgca 2582
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    434 GCCCAGTGAAGATCTGGGAGACCCTGAACTCAGAAGGCTGTGTGTCTTCTGCCCCACGCA 375

Qy   2583 cgcacccgtatctgccctccttgctggtagaagctgaagagcacggtcccccaggaggca 2642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    374 CGCACCCGTATCTGCCCTCCTTGCTGGTAGAAGCTGAAGAGCACGGTCCCCCAGGAGGCA 315

Qy   2643 gctcaggataggtggtatggagctgtgccgaggcttgggctcccacataagcactagtct 2702
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    314 GCTCAGGATAGGTGGTATGGAGCTGTGCCGAGGCTTGGGGTCCCACATAAGCACTAGTCT 255

Qy   2703 atagatgcctcttaggactggtgcctggcacagccgcgggccaggaggctgccacacgga 2762
          |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    254 ATAGATGCCTCTTAGGACTGGTGCCTGGCACAGCTGCGGGCCAGGAGGCTGCCACACGGA 195

Qy   2763 agcaagcagatgaactaatttcatttcaaggcagtttttaaagaagtcttggaaacagac 2822
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    194 AGCAAGCAGATGAACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCATGGAAACAGAC 135

Qy   2823 ggcggcacctttcctctaatccagcaaagtgattccctgcacaccagagacaagcagagt 2882
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    134 GGCGGCACCTTTCCTCTAATCCAGCAAAATGATTCCCTGCACACCAGAGACAAGCAGAGT 75

Qy   2883 aacaggatcagtgggtctaagtgtccgagacttaacgaaaatagtatttcagctgcaata 2942
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     74 AACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATA 15

Qy   2943 aagattgagtttgc 2956
          ||||||||||||||
Db     14 AAGATTGAGTTTGC 1


RESULT  52
BE243887
LOCUS       BE243887      414 bp    mRNA            EST       13-JUL-2000
DEFINITION  TCBAP1E1522 Pediatric pre-B cell acute lymphoblastic leukemia
            Baylor-HGSC project=TCBA Homo sapiens cDNA clone TCBAP1522, mRNA
            sequence.
ACCESSION   BE243887
VERSION     BE243887.1  GI:9095627
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 414)
  AUTHORS   Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman Jr.,F.R., Muzny,D.
            , Bouck,J., Gibbs,R.A. and Margolin,J.F.
  TITLE     Pediatric Leukemia cDNA Sequencing Project
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Dr. Judith F. Margolin
            Human Genome Sequencing Center at Baylor College of Medicine and
            Texas Children's Cancer Center
            One Baylor Plaza, Houston, TX 77030, USA
            Tel: 713 770 4536
            Fax: 713 770 4038
            Email: jmargolin@txccc.org
            Seq primer: M13 primer.
```

```
FEATURES             Location/Qualifiers
     source          1. .414
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TCBAP1522"
                     /clone_lib="Pediatric pre-B cell acute lymphoblastic
                     leukemia Baylor-HGSC project=TCBA"
                     /sex="male"
                     /tissue_type="leukopheresis"
                     /cell_type="pre-B cell"
                     /dev_stage="pediatric 2 years"
                     /lab_host="DH10B"
                     /note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
                     First strand cDNA was primed with an anchored
                     XhoI-oligo(dT) primer [5'GGAGGACTCGAGCGGCCGCAGGAGGAG(T)VN
                     3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second
                     strand was primed with a BamH1-dC primer
                     [5'AGAGAGCTCGGATCCGCGGCCGCAATAATAATAAT(C) 3'].
                     Double-stranded cDNA was then digested with BamH1 and
                     XhoI and directionally cloned into the BamH1 and SalI
                     sites of lambda pSB vector. Library went through one
                     round of normalization. Library was constructed by Wei
                     Yu"
BASE COUNT       72 a    125 c    143 g     74 t
ORIGIN
```

```
  Query Match             9.7%;  Score 287;  DB 105;  Length 414;
  Best Local Similarity   100.0%;  Pred. No. 2e-140;
  Matches  287;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    103 agggacgcaccatatcgcaggcacccgcccgccgcgagcggccgcgcaaggacccgctgc 162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    128 AGGGACGCACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGC 187

Qy    163 ggcacctgcgcacgcgagagaagcgcggaccgtcggggtgctccggcggcccaaacaccg 222
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    188 GGCACCTGCGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGGCCCAAACACCG 247

Qy    223 tgtacctgcaggtggtggcagcgggtagccgggactcgggcgccgcgctctacgtcttct 282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    248 TGTACCTGCAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCT 307

Qy    283 ccgagttcaaccggtatctcttcaactgtggagaaggcgttcagagactcatgcaggagc 342
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    308 CCGAGTTCAACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGC 367

Qy    343 acaagttaaaggttgctcgcctggacaacatattcctgacacgaatg 389
          |||||||||||||||||||||||||||||||||||||||||||||
Db    368 ACAAGTTAAAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATG 414
```

```
RESULT  53
AA353573
LOCUS       AA353573      355 bp    mRNA            EST       21-APR-1997
DEFINITION  EST61719 Activated T-cells XX Homo sapiens cDNA 5' end, mRNA
            sequence.
ACCESSION   AA353573
VERSION     AA353573.1  GI:2005893
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 355)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
            ,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
            ,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
            Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
            ,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,
            Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M.,
            Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.
            , Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
            Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L.,
            Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
            ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
            Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT     Other_ESTs: THC175624
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .355
                     /organism="Homo sapiens"
                     /db_xref="ATCC (inhost):153799"
                     /db_xref="taxon:9606"
                     /clone_lib="Activated T-cells XX"
                     /cell_type="T-lymphocyte"
                     /dev_stage="adult"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT       81 a    102 c     90 g     77 t      5 others
ORIGIN
```

```
  Query Match             9.7%;  Score 286;  DB 6;  Length 355;
  Best Local Similarity   100.0%;  Pred. No. 6.7e-140;
  Matches  286;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1043 aaaggcagatgcccccgtggccttggtggttcacatggccccagcatctgtgcttgtgga 1102
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AAAGGCAGATGCCCCCGTGGCCTTGGTGGTTCACATGGCCCCAGCATCTGTGCTTGTGGA 60

Qy   1103 cagcaggtaccagcagtggatggagaggtttgggcctgacacccagcacttggtcctgaa 1162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAGCAGGTACCAGCAGTGGATGGAGAGGTTTGGGCCTGACACCCAGCACTTGGTCCTGAA 120

Qy   1163 tgagaactgtgcctcagttcacaaccttcgcagccacaagattcaaacccagctcaacct 1222
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGAGAACTGTGCCTCAGTTCACAACCTTCGCAGCCACAAGATTCAAACCCAGCTCAACCT 180

Qy   1223 catccacccggacatcttccccctgctcaccagtttccgctgtaagaaggagggccccac 1282
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CATCCACCCGGACATCTTCCCCCTGCTCACCAGTTTCCGCTGTAAGAAGGAGGGCCCCAC 240

Qy   1283 cctcagtgtgcccatggttcagggtgaatgcctcctcaagtaccag 1328
          |||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTCAGTGTGCCCATGGTTCAGGGTGAATGCCTCCTCAAGTACCAG 286
```

```
RESULT  54
W37486
LOCUS       W37486        397 bp    mRNA            EST       10-OCT-1996
DEFINITION  zc10f03.s1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
            IMAGE:321917 3' similar to SW:YK59_YEAST P36159 HYPOTHETICAL 96.8
            KD PROTEIN IN SIS2-MTD1 INTERGENIC REGION. [1] ;, mRNA sequence.
ACCESSION   W37486
VERSION     W37486.1  GI:1319080
KEYWORDS    EST.
```

```
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 397)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: similarity on wrong strand
            Possible reversed clone: polyT not found
            Insert Length: 678   Std Error: 0.00
            Seq primer: mob.REGA+ET.
FEATURES             Location/Qualifiers
     source          1. .397
                     /organism="Homo sapiens"
                     /db_xref="GDB:1259575"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:321917"
                     /clone_lib="Soares_parathyroid_tumor_NbHPA"
                     /tissue_type="parathyroid tumor"
                     /dev_stage="adult"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia
                     ) with a modified polylinker; Site_1: Not I; Site_2: Eco
                     RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                     primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCACCAATTTTTTTTTTTTTTTTTT
                     TTTTT-3'], double-stranded cDNA was size selected, ligated
                     to Eco RI adapters (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of a modified pT7T3
                     vector (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
                     adenomas was kindly provided by Dr. Stephen Marx, National
                     Institute of Diabetes and Digestive and Kidney Diseases,
                     NIH."
BASE COUNT       97 a    101 c    112 g     83 t      4 others
ORIGIN
```

```
  Query Match             9.6%;  Score 285;  DB 146;  Length 397;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-139;
  Matches  285;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     252 cgggactcgggcgccgcgctctacgtcttctccgagttcaaccggtatctcttcaactgt 311
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     112 CGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTCAACCGGTATCTCTTCAACTGT 171

Qy     312 ggagaaggcgttcagagactcatgcaggagcacaagttaaaggttgctcgcctggacaac 371
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     172 GGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTAAAGGTTGCTCGCCTGGACAAC 231

Qy     372 atattcctgacacgaatgcactggtctaatgttgggggcttaagtggaatgattcttact 431
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     232 ATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGCTTAAGTGGAATGATTCTTACT 291

Qy     432 ttaaaggaaaccgggcttccaaagtgtgtactttctggacctccacaactggaaaaatac 491
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     292 TTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATAC 351

Qy     492 ctcgaagcaatcaaaatattttctggtccattgaaaggaatagaa 536
           |||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     352 CTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGAATAGAA 396


RESULT  55
AW248468/c
LOCUS       AW248468      394 bp    mRNA            EST       07-JAN-2000
DEFINITION  2820640.3prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2820640 3',
            mRNA sequence.
ACCESSION   AW248468
VERSION     AW248468.1  GI:6591461
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 394)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: 2820640.5prime
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling
            Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
            Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
            project Clone distribution: MGC clone distribution information can
            be found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality
            Scores: PHRED from University of Washington Genome Center. Vector
            Trimming: cross_match from University of Washington Genome Center
            PHRAP suite. Poly-T Identification: patMatch.pl from Berkeley
            Drosophila Genome Project. University of Washington Genome Center:
            http://www.genome.washington.edu Polyadenylation: Based upon the
            presence of a XhoI site followed by a run of 14 or more T residues
            at the beginning of the sequence, this cDNA insert was
            polyadenylated.
            Plate: LLCM4  row: L  column: 17
            High quality sequence stop: 213.
FEATURES             Location/Qualifiers
     source          1. .394
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2820640"
                     /clone_lib="NIH_MGC_7"
                     /tissue_type="small cell carcinoma"
                     /cell_line="MGC3"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT       80 a    103 c     97 g    114 t
ORIGIN
```

```
  Query Match             9.6%;  Score 284;  DB 88;  Length 394;
  Best Local Similarity   99.5%;  Pred. No. 7.7e-139;
  Matches  384;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    2573 gccccacgcacgcacccgtatctgcccctccttgctggtagaagctgaagagcacggtccc 2632
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     394 GCCCCACGCACGCACCCGTATCTGCCCTCCTTGCTGGTAGAAGCTGAAGAGCACGGTCCC 335

Qy    2633 ccaggaggcagctcaggataggtggtatggagctgtgccgaggcttgggctcccacataa 2692
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     334 CCAGGAGGCAGCTCAGGATAGGTGGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAA 275

Qy    2693 gcactagtctatagatgcctcttaggactggtgcctggcacagccgcgggccaggaggct 2752
```

```
Qy    274 GCACTAGTCTATAGATGCCTCTTAGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGTT 215

Qy   2753 gccacacggaagcaagcagatgaactaatttcatttcaaggcagtttttaaagaagtctt 2812

Db    214 GCCACACGGAAGCAAGCAGATGAACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCAT 155

Qy   2813 ggaaacagacggcggcacctttcctctaatccagcaaagtgattccctgcacaccagaga 2872

Db    154 GGAAACAGACGGCGGCACCTTTCCTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGA 95

Qy   2873 caagcagagtaacaggatcagtgggtctaagtgtccgagacttaacgaaaatagtatttc 2932

Db     94 CAAGCAGAGTAACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTC 35

Qy   2933 agctgcaataaagattgagtttgcaa 2958

Db     34 AGCTGCAATAAAGATTGAGTTTGCAA 9


RESULT  56
AI803400/c
LOCUS       AI803400      489 bp    mRNA            EST       13-DEC-1999
DEFINITION  tc42f03.x1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
            IMAGE:2067293 3', mRNA sequence.
ACCESSION   AI803400
VERSION     AI803400.1  GI:5368962
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 489)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1275   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 446.
FEATURES             Location/Qualifiers
     source          1. .489
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2067293"
                     /clone_lib="Soares_total_fetus_Nb2HF8_9w"
                     /dev_stage="8-9 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from pooled 8-9 week
                     (total) fetus material with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCTTAATTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       94 a    134 c    124 g    137 t
ORIGIN


  Query Match             9.6%;  Score 283;  DB 25;  Length 489;
  Best Local Similarity   99.2%;  Pred. No. 2.6e-138;
  Matches  483;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy   2472 cctcagcagaagcgggcccacacagaggagccacaggccaagaaggtcagagcccagtga 2531

Db    489 CCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGCCAAGAAGGTCAGAGCCCAGTGA 430

Qy   2532 agatctgggagaccctgaactcagaaggctgtgtgtcttctgccccacgcacgcacccgt 2591

Db    429 AGATCTGGGAGACCCTGAACTCAGAAGGCTGTGTGTCTTCTGCCCCACGCACGCACCCGT 370

Qy   2592 atctgccctccttgctggtagaagctgaagagcacggtcccccaggaggcagctcaggat 2651

Db    369 ATCTGCCCTCCTTGCTGGTAGAAGCTGAAGAGCACGGTCCCCCAGGAGGCAGCTCAGGAT 310

Qy   2652 aggtggtatggagctgtgccgaggcttgggctcccacataagcactagtctatagatgcc 2711

Db    309 AGGTGGTATGGAGCTGTGCCGAGGCTTGGGGTCCCACATAAGCACTAGTCTATAGATGCC 250

Qy   2712 tcttaggactggtgcctggcacagccgcgggccaggaggctgccacacggaagcaagcag 2771

Db    249 TCTTAGGACTGGTGCCTGGCACAGCTGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAG 190

Qy   2772 atgaactaatttcatttcaaggcagtttttaaagaagtcttggaaacagacggcggcacc 2831

Db    189 ATGAACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCATGGAAACAGACGGCGGCACC 130

Qy   2832 tttcctctaatccagcaaagtgattccctgcacaccagagacaagcagagtaacaggatc 2891

Db    129 TTTCCTCTAATCCAGCAAAATGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATC 70

Qy   2892 agtgggtctaagtgtccgagacttaacgaaaatagtatttcagctgcaataaagattgag 2951

Db     69 AGTGGGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAG 10

Qy   2952 tttgcaa 2958

Db      9 TTTGCAA 3


RESULT  57
T34216
LOCUS       T34216        307 bp    mRNA            EST       06-SEP-1995
DEFINITION  EST64346 Human White blood cells Homo sapiens cDNA 5' end similar
            to None, mRNA sequence.
ACCESSION   T34216
VERSION     T34216.1  GI:616314
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 307)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
            ,C.J., Lee,N., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
            ,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W., Clayton,R.A.,
            Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D., FitzGerald
            ,L.M., FitzHugh,W.M., Fritchman,J.L., Geoghagen,N.S.M., Glodek,A.,
            Gnehm,C.L., Hanna,M.C., Hedblom,E., HinkleJr,P.S., Kelley,J.M.,
            Klimek,K.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pellegrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.-J.,
            Dimke,D., Feng,P., Ferrie,A., Fischer,C., Hastings,G.A., He,W.-W.,
            Hu,J.-S., Greene,J.M., Gruber,J., Hudson,P., Kim,A., Kozak,D.L.,
            Kunsch,C., Ji,H., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wei
            ,Y.-F., Wing,J., Xu,C., Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon
            ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
            Venter,J.C.
  TITLE     Initial Assessment of Human Gene Diversity and Expression Patterns
            Based Upon 83 Million Basepairs of cDNA Sequence
  JOURNAL   Nature 377, 3-174 (1995)
  MEDLINE   96026280
COMMENT     Other_ESTs: THC15444
            Contact: Venter, JC
            The Institute for Genomic Research
            932 Clopper Rd, Gaithersburg, MD 20878
```

```
            Tel: 3018699056
            Fax: 3018699423
            Email: tdbinfo@tdb.tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please contact the TIGR Database
            (tdbinfo@tdb.tigr.org)
            Seq primer: M13 Reverse.
FEATURES              Location/Qualifiers
     source           1. .307
                      /organism="Homo sapiens"
                      /db_xref="ATCC (inhost):104947"
                      /db_xref="taxon:9606"
                      /clone_lib="Human White blood cells"
                      /tissue_type="white blood cells"
                      /note="Organ: blood"
BASE COUNT       65 a     81 c     86 g     75 t
ORIGIN


  Query Match             9.5%;  Score 282;  DB 145;  Length 307;
  Best Local Similarity   100.0%;  Pred. No. 8.7e-138;
  Matches  282;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1735 gagaacgcgccttggcatctttgggaaagccgcttcacccttgctggtggttgccccca 1794
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCCCCCA 60

Qy   1795 accagctcaaagcctggctccagcagtaccacaaccagtgccaggaggtcctgcaccaca 1854
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCACCACA 120

Qy   1855 tcagtatgattcctgccaaatgccttcaggaaggggctgagatctccagtcctgcagtgg 1914
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTCCTGCAGTGG 180

Qy   1915 aaagattgatcagttcgctgttgcgaacatgtgatttggaagagtttcagacctgtctgg 1974
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGACCTGTCTGG 240

Qy   1975 tgcggcactgcaagcatgcgtttggctgtgcgctggtgcaca 2016
          |||||||||||||||||||||||||||||||||||||||||
Db    241 TGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACA 282


RESULT  58
AA534478/c
LOCUS       AA534478      489 bp    mRNA            EST       21-AUG-1997
DEFINITION  nf76f10.s1 NCI_CGAP_Co3 Homo sapiens cDNA clone IMAGE:925867 3',
            mRNA sequence.
ACCESSION   AA534478
VERSION     AA534478.1  GI:2278731
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 489)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D.
            , Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arraying: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 680    Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 336.
FEATURES              Location/Qualifiers
     source           1. .489
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:925867"
                      /clone_lib="NCI_CGAP_Co3"
                      /sex="pooled"
                      /tissue_type="colon"
                      /lab_host="DH10B"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was prepared from 12 pooled bulk tumor samples and primed
                      with a Not I - oligo(dT) primer. Double-stranded cDNA was
                      ligated to Eco RI adaptors (Pharmacia), digested with Not
                      I and cloned into the Not I and Eco RI sites of the
                      modified pT7T3 vector. Library went through one round of
                      normalization. "
BASE COUNT       95 a    135 c    123 g    136 t
ORIGIN


  Query Match             9.5%;  Score 281;  DB 8;  Length 489;
  Best Local Similarity   99.2%;  Pred. No. 3e-137;
  Matches  481;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy   2474 tcagcagaagcgggcccacacagaggagccacaggccaagaaggtcagagcccagtgaag 2533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    485 TCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGCCAAGAAGGTCAGAGCCCAGTGAAG 426

Qy   2534 atctgggagaccctgaactcagaaggctgtgtgtcttctgccccacgcacgcacccgtat 2593
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    425 ATCTGGGAGACCCTGAACTCAGAAGGCTGTGTGTCTTCTGCCCCACGCACGCACCCGTAT 366

Qy   2594 ctgccctccttgctggtagaagctgaagagcacggtcccccaggaggcagctcaggatag 2653
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    365 CTGCCCTCCTTGCTGGTAGAAGCTGAAGAGCACGGTCCCCCAGGAGGCAGCTCAGGATAG 306

Qy   2654 gtggtatggagctgtgccgaggcttgggctcccacataagcactagtctatagatgcctc 2713
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    305 GTGGTATGGAGCTGTGCCGAGGCTTGGGGTCCCACATAAGCACTAGTCTATAGATGCCTC 246

Qy   2714 ttaggactggtgcctggcacagccgcgggccaggaggctgccacacggaagcaagcagat 2773
          |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    245 TTAGGACTGGTGCCTGGCACAGCTGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGAT 186

Qy   2774 gaactaatttcatttcaaggcagtttttaaagaagtcttggaaacagacggcggcacctt 2833
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    185 GAACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCATGGAAACAGACGGCGGCACCTT 126

Qy   2834 tcctctaatccagcaaagtgattccctgcacaccagagacaagcagagtaacaggatcag 2893
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    125 TCCTCTAATCCAGCAAAATGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAG 66

Qy   2894 tgggtctaagtgtccgagacttaacgaaaatagtatttcagctgcaataaagattgagtt 2953
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     65 TGGGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTT 6

Qy   2954 tgcaa 2958
          |||||
Db      5 TGCAA 1


RESULT  59
AW806551
LOCUS       AW806551      278 bp    mRNA            EST       17-MAY-2000
DEFINITION  IL0-ST0002-160599-003 ST0002 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW806551
VERSION     AW806551.1  GI:7899550
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
EFERENCE    1  (bases 1 to 278)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
OMMENT      Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=IL0-ST0002-160
            599-003&t3=1999-05-16&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 278.
EATURES             Location/Qualifiers
     source         1. .278
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone_lib="ST0002"
                    /dev_stage="Adult"
                    /note="Organ: stomach; Vector: puc18; Site_1: SmaI;
                    Site_2: SmaI; A mini-library was made by cloning products
                    derived from  ORESTES PCR (U.S. Letters Patent application
                    No. 196,716 -  Ludwig Institute for Cancer Research)
                    profiles into  the pUC 18 vector. Reverse transcription of
                    tissue mRNA and cDNA amplification were performed under
                    low stringency conditions."
ASE COUNT       64 a     84 c     70 g     60 t
RIGIN


  Query Match             9.3%;  Score 276;  DB 95;  Length 278;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-134;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1123 tggagaggtttgggcctgacacccagcacttggtcctgaatgagaactgtgcctcagttc 1182
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3 TGGAGAGGTTTGGGCCTGACACCCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTC 62

Qy   1183 acaaccttcgcagccacaagattcaaacccagctcaacctcatccacccggacatcttcc 1242
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     63 ACAACCTTCGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCC 122

Qy   1243 ccctgctcaccagtttccgctgtaagaaggagggccccaccctcagtgtgcccatggttc 1302
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 CCCTGCTCACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTC 182

Qy   1303 agggtgaatgcctcctcaagtaccagctccgtcccaggagggagtggcagagggatgcca 1362
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 AGGGTGAATGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCA 242

Qy   1363 ttattacttgcaatcctgaggaattcatagttgagg 1398
            |||||||||||||||||||||||||||||||||||
Db    243 TTATTACTTGCAATCCTGAGGAATTCATAGTTGAGG 278


ESULT  60
A635046
OCUS        AA635046      461 bp    mRNA            EST       06-MAR-1998
EFINITION   ab48b06.r1 Stratagene lung carcinoma 937218 Homo sapiens cDNA clone
            IMAGE:844019 5', mRNA sequence.
ACCESSION   AA635046
VERSION     AA635046.1  GI:2558260
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 461)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 694    Std Error: 0.00
            Seq primer: -28m13 revl ET from Amersham
            High quality sequence stop: 453.
FEATURES            Location/Qualifiers
     source         1. .461
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:844019"
                    /clone_lib="Stratagene lung carcinoma 937218"
                    /tissue_type="lung carcinoma"
                    /cell_line="NCI-H69"
                    /dev_stage="cell line NCI-H69"
                    /lab_host="SOLR (kanamycin resistant)"
                    /note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI
                    ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                    dT. Small cell carcinoma cell line NCI-H69.  Average
                    insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor
                    sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
                    CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT      110 a    120 c    159 g     72 t
ORIGIN


  Query Match             9.3%;  Score 276;  DB 10;  Length 461;
  Best Local Similarity   99.7%;  Pred. No. 1.3e-134;
  Matches  326;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   2410 tgcggcaggtgcgggcggccctcctgtccagggagctggcaggcggcctggaggatgggg 2469
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     42 TGCGGCAGGTGCGGGCGGCCCTCCTGTCCAGGGAGCTGGCAGGCGGCCTGGAGGATGGGG 101

Qy   2470 agcctcagcagaagcgggcccacacagaggagccacaggccaagaaggtcagagcccagt 2529
            |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db    102 AGCCTCAGCAGAAGCGGGCCCACACAGAGGAGCCACAGGCGAAGAAGGTCAGAGCCCAGT 161

Qy   2530 gaagatctgggagaccctgaactcagaaggctgtgtgtcttctgccccacgcacgcaccc 2589
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    162 GAAGATCTGGGAGACCCTGAACTCAGAAGGCTGTGTGTCTTCTGCCCCACGCACGCACCC 221

Qy   2590 gtatctgccctccttgctggtagaagctgaagagcacggtcccccaggaggcagctcagg 2649
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    222 GTATCTGCCCTCCTTGCTGGTAGAAGCTGAAGAGCACGGTCCCCCAGGAGGCAGCTCAGG 281

Qy   2650 ataggtggtatggagctgtgccgaggcttgggctcccacataagcactagtctatagatg 2709
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    282 ATAGGTGGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATG 341

Qy   2710 cctcttaggactggtgcctggcacagc 2736
            |||||||||||||||||||||||||||
Db    342 CCTCTTAGGACTGGTGCCTGGCACAGC 368
```

```
RESULT  61
AW175581/c
LOCUS       AW175581      364 bp    mRNA            EST       16-NOV-1999
DEFINITION  QV0-BT0041-030999-013-e09 BT0041 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW175581
VERSION     AW175581.1  GI:6441618
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 364)
  AUTHORS   HCGP http://www.ludwig.org.br/ORESTES.
  TITLE     The FAPESP/LICR Human Cancer Genome Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV0&t2=QV0-BT0041-
            030999-013-e09&t3=1999-09-03&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 75
            High quality sequence stop: 364.
FEATURES             Location/Qualifiers
     source          1. .364
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="BT0041"
                     /dev_stage="Adult"
                     /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No. 196
                     ,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
BASE COUNT       75 a     93 c    118 g     78 t
ORIGIN
```

```
  Query Match             9.3%;  Score 274;  DB 87;  Length 364;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-133;
  Matches  274;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1025 ctttcagaggtaccaaggaaaggcagatgccccgtggccttggtggttcacatggcccc 1084
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     364 CTTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTGGTTCACATGGCCCC 305

Qy    1085 agcatctgtgcttgtggacagcaggtaccagcagtggatggagaggtttgggcctgacac 1144
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     304 AGCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGGTTTGGGCCTGACAC 245

Qy    1145 ccagcacttggtcctgaatgagaactgtgcctcagttcacaaccttcgcagccacaagat 1204
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     244 CCAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTTCGCAGCCACAAGAT 185

Qy    1205 tcaaacccagctcaacctcatccacccggacatcttcccccctgctcaccagtttccgctg 1264
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     184 TCAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTCACCAGTTTCCGCTG 125

Qy    1265 taagaaggagggccccacccctcagtgtgcccatg 1298
           ||||||||||||||||||||||||||||||||||
Db     124 TAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATG 91
```

```
RESULT  62
AA994126
LOCUS       AA994126      457 bp    mRNA            EST       27-AUG-1998
DEFINITION  ou38b06.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:1628531 3' similar to SW:YK59_YEAST P36159 HYPOTHETICAL 96.8
            KD PROTEIN IN SIS2-MTD1 INTERGENIC REGION. ;, mRNA sequence.
ACCESSION   AA994126
VERSION     AA994126.1  GI:3180671
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 457)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1193   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 439.
FEATURES             Location/Qualifiers
     source          1. .457
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1628531"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "
BASE COUNT      115 a    119 c    127 g     96 t
ORIGIN
```

```
  Query Match             9.2%;  Score 273;  DB 14;  Length 457;
  Best Local Similarity   100.0%;  Pred. No. 5.2e-133;
  Matches  273;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     279 ttctccgagttcaaccggtatctcttcaactgtggagaaggcgttcagagactcatgcag 338
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     138 TTCTCCGAGTTCAACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAG 197

Qy     339 gagcacaagttaaaggttgctcgcctggacaacatattcctgacacgaatgcactggtct 398
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     198 GAGCACAAGTTAAAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCT 257

Qy     399 aatgttgggggcttaagtggaatgattcttactttaaaggaaaccgggcttccaaagtgt 458
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     258 AATGTTGGGGGCTTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGT 317

Qy     459 gtactttctggacctccacaactggaaaaatacctcgaagcaatcaaaatattttctggt 518
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     318 GTACTTTCTGGACCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGT 377

Qy     519 ccattgaaaggaatagaactggctgtgcggccc 551
           |||||||||||||||||||||||||||||||
Db     378 CCATTGAAAGGAATAGAACTGGCTGTGCGGCCC 410
```

```
RESULT  63
BE744876
LOCUS       BE744876      984 bp    mRNA            EST       15-SEP-2000
DEFINITION  601576324F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3837533 5',
            mRNA sequence.
ACCESSION   BE744876
VERSION     BE744876.1  GI:10158868
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 984)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM522  row: o  column: 06
            High quality sequence stop: 726.
FEATURES             Location/Qualifiers
     source          1. .984
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3837533"
                     /clone_lib="NIH_MGC_9"
                     /tissue_type="adenocarcinoma cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) and Superscript II RT (Life
                     Technologies)."
BASE COUNT      227 a    259 c    316 g    182 t
ORIGIN
```

```
  Query Match             9.0%;  Score 266;  DB 135;  Length 984;
  Best Local Similarity   99.5%;  Pred. No. 2.8e-129;
  Matches  436;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy    966 tttgtggtggtagaatgtccagatgaaagcttcattcaacccatctgtgagaatgccacc 1025
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 TTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATCTGTGAGAATGCCACC 303

Qy   1026 tttcagaggtaccaaggaaaggcagatgcccccgtggccttggtggttcacatggcccca 1085
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 TTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTGGTTCACATGGCCCCA 363

Qy   1086 gcatctgtgcttgtggacagcaggtaccagcagtggatggagaggtttgggcctgacacc 1145
          |||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    364 GCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGG-TTGGGCCTGACACC 422

Qy   1146 cagcacttggtcctgaatgagaactgtgcctcagttcacaaccttcgcagccacaagatt 1205
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    423 CAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTTCGCAGCCACAAGATT 482

Qy   1206 caaacccagctcaacctcatccacccggacatcttccccctgctcaccagtttccgctgt 1265
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    483 CAAACCCAGCTCAACCTCATCCACCCGGACATCTTCCCCCTGCTCACCAGTTTCCGCTGT 542

Qy   1266 aagaaggagggccccaccctcagtgtgcccatggttcagggtgaatgcctcctcaagtac 1325
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    543 AAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGGTTCAGGGTGAATGGCTCCTCAAGTAC 602

Qy   1326 cagctccgtcccaggagggagtggcagagggatgccattattacttgcaatcctgaggaa 1385
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    603 CAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATGCCATTATTACTTGCAATCCTGAGGAA 662

Qy   1386 ttcatagttgaggcgctg 1403
          ||||||||||||||||||
Db    663 TTCATAGTTGAGGCGCTG 680
```

```
RESULT  64
AI033108/c
LOCUS       AI033108      410 bp    mRNA            EST       28-AUG-1998
DEFINITION  ow98g08.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA
            clone IMAGE:1654910 3', mRNA sequence.
ACCESSION   AI033108
VERSION     AI033108.1  GI:3254061
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 410)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 773   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 368.
FEATURES             Location/Qualifiers
     source          1. .410
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1654910"
                     /clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     This is a subtracted version of the original Soares fetal
                     liver spleen 1NFLS library.   1st strand cDNA was primed
                     with a Pac I - oligo(dT) primer [5'
                     AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       88 a    110 c    101 g    111 t
ORIGIN
```

```
  Query Match             8.6%;  Score 255;  DB 15;  Length 410;
  Best Local Similarity   99.3%;  Pred. No. 1.8e-123;
  Matches  405;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy   2551 ctcagaaggctgtgtgtcttctgccccacgcacgcacccgtatctgccctccttgctggt 2610
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    410 CTCAGAAGGCTGTGTGTCTTCTGCCCCACGCACGCACCCGTATCTGCCCTCCTTGCTGGT 351

Qy   2611 agaagctgaagagcacggtcccccaggaggcagctcaggataggtggtatggagctgtgc 2670
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    350 AGAAGCTGAAGAGCACGGTCCCCCAGGAGGCAGCTCAGGATAGGTGGTATGGAGCTGTGC 291
```

```
Qy   2671 cgaggcttgggctcccacataagcactagtctatagatgcctcttaggactggtgcctgg 2730
          |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    290 CGAGGCTTGGGGTCCCACATAAGCACTAGTCTATAGATGCCTCTTAGGACTGGTGCCTGG 231

Qy   2731 cacagccgcgggccaggaggctgccacacggaagcaagcagatgaactaatttcatttca 2790
          |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    230 CACAGCTGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGAACTAATTTCATTTCA 171

Qy   2791 aggcagtttttaaagaagtcttggaaacagacggcggcacctttcctctaatccagcaaa 2850
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    170 AGGCAGTTTTTAAAGAAGTCATGGAAACAGACGGCGGCACCTTTCCTCTAATCCAGCAAA 111

Qy   2851 gtgattccctgcacaccagagacaagcagagtaacaggatcagtgggtctaagtgtccga 2910
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    110 GTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTGGGTCTAAGTGTCCGA 51

Qy   2911 gacttaacgaaaatagtatttcagctgcaataaagattgagtttgcaa 2958
          |||||||||||||||||||||||||||||||||||||||||||||||
Db     50 GACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTGCAA 3
```

```
RESULT  65
AW592223/c
LOCUS       AW592223      461 bp    mRNA            EST       22-MAR-2000
DEFINITION  hf41a01.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:2934408 3', mRNA sequence.
ACCESSION   AW592223
VERSION     AW592223.1  GI:7279399
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 461)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 450.
FEATURES             Location/Qualifiers
     source          1. .461
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2934408"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo. "
BASE COUNT       93 a    125 c    114 g    129 t
ORIGIN
```

```
  Query Match             8.6%;  Score 255;  DB 92;  Length 461;
  Best Local Similarity   99.1%;  Pred. No. 1.8e-123;
  Matches  455;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy   2500 agccacaggccaagaaggtcagagcccagtgaagatctgggagaccctgaactcagaagg 2559
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    461 AGCCACAGGCCAAGAAGGTCAGAGCCCAGTGAAGATCTGGGAGACCCTGAACTCAGAAGG 402

Qy   2560 ctgtgtgtcttctgccccacgcacgcacccgtatctgccctccttgctggtagaagctga 2619
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    401 CTGTGTGTCTTCTGCCCCACGCACGCACCCGTATCTGCCCTCCTTGCTGGTAGAAGCTGA 342

Qy   2620 agagcacggtccccccaggaggcagctcaggataggtggtatggagctgtgccgaggcttg 2679
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    341 AGAGCACGGTCCCCCAGGAGGCAGCTCAGGATAGGTGGTATGGAGCTGTGCCGAGGCTTG 282

Qy   2680 ggctcccacataagcactagtctatagatgcctcttaggactggtgcctggcacagccgc 2739
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    281 GGGTCCCACATAAGCACTAGTCTATAGATGCCTCTTAGGACTGGTGCCTGGCACAGCTGC 222

Qy   2740 gggccaggaggctgccacacggaagcaagcagatgaactaatttcatttcaaggcagttt 2799
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 GGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGAACTAATTTCATTTCAAGGCAGTTT 162

Qy   2800 ttaaagaagtcttggaaacagacggcggcacctttcctctaatccagcaaagtgattccc 2859
          |||||||||||| ||||||||||||||||||||||||||||||||||||| ||||||||
Db    161 TTAAAGAAGTCATGGAAACAGACGGCGGCACCTTTCCTCTAATCCAGCAAAATGATTCCC 102

Qy   2860 tgcacaccagagacaagcagagtaacaggatcagtgggtctaagtgtccgagacttaacg 2919
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 TGCACACCAGAGACAAGCAGAGTAACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACG 42

Qy   2920 aaaatagtatttcagctgcaataaagattgagtttgcaa 2958
          |||||||||||||||||||||||||||||||||||||
Db     41 AAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTGCAA 3
```

```
RESULT  66
AA716607/c
LOCUS       AA716607      479 bp    mRNA            EST       29-DEC-1997
DEFINITION  zg68g07.s1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone
            IMAGE:398556 3', mRNA sequence.
ACCESSION   AA716607
VERSION     AA716607.1  GI:2728881
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 479)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 462.
FEATURES             Location/Qualifiers
     source          1. .479
                     /organism="Homo sapiens"
                     /db_xref="GDB:1306379"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:398556"
                     /clone_lib="Soares_fetal_heart_NbHH19W"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
```

modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W."

ASE COUNT 93 a 127 c 118 g 141 t
RIGIN

Query Match 8.6%; Score 253; DB 11; Length 479;
Best Local Similarity 99.1%; Pred. No. 2.1e-122;
Matches 453; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
y  2502 ccacaggccaagaaggtcagagcccagtgaagatctgggagaccctgaactcagaaggct 2561
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   467 CCACAGGCCAAGAAGGTCAGAGCCCAGTGAAGATCTGGGAGACCCTGAACTCAGAAGGCT 408

y  2562 gtgtgtcttctgccccacgcacgcacccgtatctgccctccttgctggtagaagctgaag 2621
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   407 GTGTGTCTTCTGCCCCACGCACGCACCCGTATCTGCCCTCCTTGCTGGTAGAAGCTGAAG 348

y  2622 agcacggtcccccaggaggcagctcaggataggtggtatggagctgtgccgaggcttggg 2681
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   347 AGCACGGTCCCCCAGGAGGCAGCTCAGGATAGGTGGTATGGAGCTGTGCCGAGGCTTGGG 288

y  2682 ctcccacataagcactagtctatagatgcctcttaggactggtgcctggcacagccgcgg 2741
        |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
b   287 GTCCCACATAAGCACTAGTCTATAGATGCCTCTTAGGACTGGTGCCTGGCACAGCTGCGG 228

y  2742 gccaggaggctgccacacggaagcaagcagatgaactaatttcatttcaaggcagttttt 2801
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   227 GCCAGGAGGCTGCCACACGGAAGCAAGCAGATGAACTAATTTCATTTCAAGGCAGTTTTT 168

y  2802 aaagaagtcttggaaacagacggcggcacctttcctctaatccagcaaagtgattccctg 2861
        ||||||||| ||||||||||||||||||||||||||||||||||||| ||||||||||||
b   167 AAAGAAGTCATGGAAACAGACGGCGGCACCTTTCCTCTAATCCAGCAAAATGATTCCCTG 108

y  2862 cacaccagagacaagcagagtaacaggatcagtgggtctaagtgtccgagacttaacgaa 2921
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b   107 CACACCAGAGACAAGCAGAGTAACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACGAA 48

y  2922 aatagtatttcagctgcaataaagattgagtttgcaa 2958
        |||||||||||||||||||||||||||||||||||
b    47 AATAGTATTTCAGCTGCAATAAAGATTGAGTTTGCAA 11
```

ESULT 67
A311855
OCUS AA311855 501 bp mRNA EST 19-APR-1997
EFINITION EST182568 Jurkat T-cells VI Homo sapiens cDNA 5' end, mRNA sequence.
CCESSION AA311855
ERSION AA311855.1 GI:1964184
EYWORDS EST.
OURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
EFERENCE 1 (bases 1 to 501)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.
TITLE Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
COMMENT Other_ESTs: THC175624
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.
FEATURES Location/Qualifiers
source 1. .501
/organism="Homo sapiens"
/db_xref="ATCC (inhost):159063"
/db_xref="taxon:9606"
/clone_lib="Jurkat T-cells VI"
/cell_type="T-lymphocyte"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 125 a 148 c 125 g 98 t 5 others
ORIGIN

Query Match 8.5%; Score 250; DB 5; Length 501;
Best Local Similarity 100.0%; Pred. No. 8.1e-121;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy  1180 ttcacaaccttcgcagccacaagattcaaacccagctcaacctcatccacccggacatct 1239
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   100 TTCACAACCTTCGCAGCCACAAGATTCAAACCCAGCTCAACCTCATCCACCCGGACATCT 159

Qy  1240 tccccctgctcaccagtttccgctgtaagaaggagggccccacccctcagtgtgcccatgg 1299
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   160 TCCCCCTGCTCACCAGTTTCCGCTGTAAGAAGGAGGGCCCCACCCTCAGTGTGCCCATGG 219

Qy  1300 ttcagggtgaatgcctcctcaagtaccagctccgtcccaggagggagtggcagagggatg 1359
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   220 TTCAGGGTGAATGCCTCCTCAAGTACCAGCTCCGTCCCAGGAGGGAGTGGCAGAGGGATG 279

Qy  1360 ccattattacttgcaatcctgaggaattcatagttgaggcgctgcagcttcccaacttcc 1419
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   280 CCATTATTACTTGCAATCCTGAGGAATTCATAGTTGAGGCGCTGCAGCTTCCCAACTTCC 339

Qy  1420 agcagagcgt 1429
         ||||||||||
Db   340 AGCAGAGCGT 349
```

RESULT 68
Z17886
LOCUS Z17886 299 bp mRNA EST 10-NOV-1992
DEFINITION HSDHEI030 Stratagene cDNA library Human heart, cat#936208 Homo sapiens cDNA clone HEI030, mRNA sequence.
ACCESSION Z17886
VERSION Z17886.1 GI:30720
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 299)
  AUTHORS  Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes
           ,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F.,
           Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,
           Sebastiani-Kabaktchis,C. and Tessier,A.
  TITLE    IMAGE: molecular integration of the analysis of the human genome
           and its expression
  JOURNAL  C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
  MEDLINE  95277534
COMMENT    Contact: Genzentrum Muenchen
           Laboratorium fuer molekulare Biologie
           Am Klopferspitz 18a,8033 Martinsried,Germany
           Email: obermaier@vms.biochem.mpg.de
           single read.
FEATURES             Location/Qualifiers
     source          1. .299
                     /organism="Homo sapiens"
                     /db_xref="GDB:D0S7384E"
                     /db_xref="taxon:9606"
                     /clone="HEI030"
                     /clone_lib="Stratagene cDNA library Human heart,
                     cat#936208"
                     /note="Vector: pBluescript SK(+); Human heart cDNA
                     library. Cloning vector pBluescript SK(+)"
BASE COUNT       96 a     77 c     70 g     56 t
ORIGIN


  Query Match             8.4%;  Score 248;  DB 147;  Length 299;
  Best Local Similarity   99.7%;  Pred. No. 9e-120;
  Matches  298;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    451 caaagtgtgtactttctggacctccacaactggaaaaatacctcgaagcaatcaaaatat 510
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CAAAGTGTGTACTTTCTGGACCTCCACAACTGGAAAAATACCTCGAAGCAATCAAAATAT 60

Qy    511 tttctggtccattgaaaggaatagaactggctgtgcggccccactctgccccagaatacg 570
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTTCTGGTCCATTGAAAGGAATAGAACTGGCTGTGCGGCCCCACTCTGCCCCAGAATACT 120

Qy    571 aggatgaaaccatgacagtttaccagatccccatacacagtgaacagaggaggggaaagc 630
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AGGATGAAACCATGACAGTTTACCAGATCCCCATACACAGTGAACAGAGGAGGGGAAAGC 180

Qy    631 accaaccatggcagagtccagaaaggcctctcagcaggctcagtccagagcgatcttcag 690
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACCAACCATGGCAGAGTCCAGAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGATCTTCAG 240

Qy    691 actccgagtcgaatgaaaatgagccacaccttccacatggtgttagccagagaagaggg 749
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ACTCCGAGTCGAATGAAAATGAGCCACACCTTCCACATGGTGTTAGCCAGAGAAGAGGG 299


RESULT  69
AI991599/c
LOCUS      AI991599      517 bp    mRNA            EST       08-MAR-2000
DEFINITION ws18c04.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2497542 3'
           similar to SW:YATA_SCHPO Q10155 HYPOTHETICAL 90.6 KD PROTEIN
           C1D4.10 IN CHROMOSOME I. ;, mRNA sequence.
ACCESSION  AI991599
VERSION    AI991599.1  GI:5838504
KEYWORDS   EST.
SOURCE     human.
  ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 517)
  AUTHORS  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE    National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
           Tumor Gene Index
  JOURNAL  Unpublished (1997)
COMMENT    Contact: Robert Strausberg, Ph.D.
           Tel: (301) 496-1550
           Email: Robert_Strausberg@nih.gov
           Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
           R. Emmert-Buck, M.D., Ph.D.
           cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
           Bonaldo, Ph.D.
           cDNA Library Arrayed by: Greg Lennon, Ph.D.
           DNA Sequencing by: Washington University Genome Sequencing Center
           Clone distribution: NCI-CGAP clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           www-bio.llnl.gov/bbrp/image/image.html
           Insert Length: 625   Std Error: 0.00
           Seq primer: -40UP from Gibco
           High quality sequence stop: 426.
FEATURES             Location/Qualifiers
     source          1. .517
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2497542"
                     /clone_lib="NCI_CGAP_GC6"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Plasmid DNA from the normalized library
                     NCI_CGAP_GC4 was prepared, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from a pool of 5,000 clones made
                     from the same library (cloneIDs 1257096-1258631,
                     1469064-1470983, and 1475592-1476743). Subtraction by
                     Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      118 a    139 c    123 g    136 t      1 others
ORIGIN


  Query Match             8.4%;  Score 247;  DB 27;  Length 517;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-119;
  Matches  247;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1846 tgcaccacatcagtatgattcctgccaaatgccttcaggaaggggctgagatctccagtc 1905
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    447 TGCACCACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCTCCAGTC 388

Qy   1906 ctgcagtggaaagattgatcagttcgctgttgcgaacatgtgatttggaagagtttcaga 1965
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    387 CTGCAGTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGATTTGGAAGAGTTTCAGA 328

Qy   1966 cctgtctggtgcggcactgcaagcatgcgtttggctgtgcgctggtgcacacctctggct 2025
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    327 CCTGTCTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCTGGTGCACACCTCTGGCT 268

Qy   2026 ggaaagtggtctattccggggacaccatgccctgcgaggctctggtccggatggggaaag 2085
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    267 GGAAAGTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGGTCCGGATGGGGAAAG 208

Qy   2086 atgccac 2092
          |||||||
Db    207 ATGCCAC 201


RESULT  70
BE938229
LOCUS      BE938229      387 bp    mRNA            EST       02-OCT-2000
DEFINITION CM4-TN0060-290800-565-B02 TN0060 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BE938229
VERSION    BE938229.1  GI:10465340
KEYWORDS   EST.
SOURCE     human.
  ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 387)
```

```
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
OMMENT      Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM4-TN0060-290
            800-565-B02&t3=2000-08-29&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 385.
EATURES             Location/Qualifiers
     source         1. .387
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone_lib="TN0060"
                    /dev_stage="Adult"
                    /note="Organ: testis_normal; Vector: puc18; Site_1: SmaI;
                    Site_2: SmaI; A mini-library was made by cloning products
                    derived from  ORESTES PCR (U.S. Letters Patent application
                    No. 196,716 -  Ludwig Institute for Cancer Research)
                    profiles into  the pUC 18 vector. Reverse transcription of
                    tissue mRNA and cDNA amplification were performed under
                    low stringency conditions."
ASE COUNT       112 a    108 c     94 g     72 t      1 others
RIGIN


  Query Match             8.3%;  Score 246;  DB 137;  Length 387;
  Best Local Similarity   99.4%;  Pred. No. 1e-118;
  Matches  346;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

y     564 gaatacgaggatgaaaccatgacagtttaccagatccccatacacagtgaacagaggagg 623
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b      40 GAATACGAGGATGAAACCATGACAGTTTACCAGATCCCCATACACAGTGAACAGAGGAGG 99

y     624 ggaaagcaccaaccatggcagagtccagaaaggcctctcagcaggctcagtccagagcga 683
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     100 GGAAAGCACCAACCATGGCAGAGTCCAGAAAGGCCTCTCAGCAGGCTCAGTCCAGAGCGA 159

y     684 tcttcagactccgagtcgaatgaaaatgagccacaccttccacatggtgttagccagaga 743
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     160 TCTTCAGACTCCGAGTCGAATGAAAATGAGCCACACCTTCCACATGGTGTTAGCCAGAGA 219

y     744 agaggggtcagggactcttccctggtcgtagctttcatctgtaagcttcacttaaagaga 803
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     220 AGAGGGGTCAGGGACTCTTCCCTGGTCGTAGCTTTCATCTGTAAGCTTCACTTAAAGAGA 279

y     804 ggaaacttcttggtgctcaaagcaaaggagatgggcctcccagttgggacagctgccatc 863
          ||||| || ||||||||||||||||||||||||||||||||||||||||||||||||||||
b     280 GGAAATTTNTTGGTGCTCAAAGCAAAGGAGATGGGCCTCCCAGTTGGGACAGCTGCCATC 339

y     864 gctcccatcattgctgctgtcaaggacgggaaaagcatcactcatgaa 911
          ||||||||||||||||||||||||||||||||||||||||||||||||
b     340 GCTCCCATCATTGCTGCTGTCAAGGACGGGAAAAGCATCACTCATGAA 387


ESULT  71
838624/c
CUS       AA838624      429 bp    mRNA            EST       18-MAR-1998

DEFINITION  oe91f04.s1 NCI_CGAP_Co12 Homo sapiens cDNA clone IMAGE:1419007,
            mRNA sequence.
ACCESSION   AA838624
VERSION     AA838624.1  GI:2913423
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 429)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: Stratagene, Inc.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1787   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 375.
FEATURES            Location/Qualifiers
     source         1. .429
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:1419007"
                    /clone_lib="NCI_CGAP_Co12"
                    /sex="mixed"
                    /tissue_type="colon tumor"
                    /lab_host="SOLR (kanamycin resistant)"
                    /note="Organ: colon; Vector: Bluescript SK-; Site_1: EcoRI
                    ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                    dT. Pooled colon tumors. 5' adaptor sequence: 5'
                    GAATTCGGCACGAG 3' 3' adaptor sequence: 5'
                    CTCGAGTTTTTTTTTTTTTTTT 3' Average insert size:  1.2 kb."
BASE COUNT      91 a    115 c    108 g    115 t
ORIGIN


  Query Match             8.3%;  Score 245;  DB 12;  Length 429;
  Best Local Similarity   99.5%;  Pred. No. 3.6e-118;
  Matches  415;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy   2539 ggagaccctgaactcagaaggctgtgtgtcttctgccccacgcacgcacccgtatctgcc 2598
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    416 GGAGACCCTGAACTCAGAAGGCTGTGTGTCTTCTGCCCCACGCACGCACCCGTATCTGCC 357

Qy   2599 ctccttgctggtagaagctgaagagcacggtcccccaggaggcagctcaggataggtggt 2658
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    356 CTCCTTGCTGGTAGAAGCTGAAGAGCACGGTCCCCCAGGAGGCAGCTCAGGATAGGTGGT 297

Qy   2659 atggagctgtgccgaggcttgggctcccacataagcactagtctatagatgcctcttagg 2718
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    296 ATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTTAGG 237

Qy   2719 actggtgcctggcacagccgcgggccaggaggctgccacacggaagcaagcagatgaact 2778
           |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    236 ACTGGTGCCT-GCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGAACT 178

Qy   2779 aatttcatttcaaggcagtttttaaagaagtcttggaaacagacggcggcacctttcctc 2838
           |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    177 AATTTCATTTCAAGGCAGTTTTTAAAGAAGTCATGGAAACAGACGGCGGCACCTTTCCTC 118

Qy   2839 taatccagcaaagtgattccctgcacaccagagacaagcagagtaacaggatcagtgggt 2898
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    117 TAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTGGGT 58
```

```
Qy   2899 ctaagtgtccgagacttaacgaaaatagtatttcagctgcaataaagattgagtttg 2955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     57 CTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTG 1


RESULT 72
AW889463/c
LOCUS       AW889463      346 bp    mRNA            EST       24-MAY-2000
DEFINITION  RC6-NT0029-240400-011-B07 NT0029 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW889463
VERSION     AW889463.1  GI:8053668
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 346)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC6-NT0029-240
            400-011-B07&t3=2000-04-24&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 12
            High quality sequence stop: 310.
FEATURES             Location/Qualifiers
     source          1. .346
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="NT0029"
                     /dev_stage="Adult"
                     /note="Organ: nervous_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       75 a    103 c     85 g     83 t
ORIGIN


  Query Match             8.2%;  Score 243;  DB 96;  Length 346;
  Best Local Similarity  100.0%;  Pred. No. 4e-117;
  Matches  243;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1829 ccagtgccaggaggtcctgcaccacatcagtatgattcctgccaaatgccttcaggaagg 1888
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    335 CCAGTGCCAGGAGGTCCTGCACCACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGG 276

Qy   1889 ggctgagatctccagtcctgcagtggaaagattgatcagttcgctgttgcgaacatgtga 1948
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    275 GGCTGAGATCTCCAGTCCTGCAGTGGAAAGATTGATCAGTTCGCTGTTGCGAACATGTGA 216

Qy   1949 tttggaagagtttcagacctgtctggtgcggcactgcaagcatgcgtttggctgtgcgct 2008
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    215 TTTGGAAGAGTTTCAGACCTGTCTGGTGCGGCACTGCAAGCATGCGTTTGGCTGTGCGCT 156

Qy   2009 ggtgcacacctctggctggaaagtggtctattccggggacaccatgccctgcgaggctct 2068
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    155 GGTGCACACCTCTGGCTGGAAAGTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCT 96

Qy   2069 ggt 2071
          |||
Db     95 GGT 93


RESULT 73
H22087/c
LOCUS       H22087        414 bp    mRNA            EST       06-JUL-1995
DEFINITION  yl34c04.r1 Soares breast 3NbHBst Homo sapiens cDNA clone
            IMAGE:160134 5' similar to SP:YK59_YEAST P36159 HYPOTHETICAL 96.8
            KD PROTEIN IN SIS2-MTD1 INTERGENIC ;, mRNA sequence.
ACCESSION   H22087
VERSION     H22087.1  GI:890782
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 414)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 490
            High quality sequence stops: 215
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: similarity on wrong strand
            Insert Length: 490   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 215.
FEATURES             Location/Qualifiers
     source          1. .414
                     /organism="Homo sapiens"
                     /db_xref="GDB:578554"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:160134"
                     /clone_lib="Soares breast 3NbHBst"
                     /sex="Female"
                     /dev_stage="adult"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: breast; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of a modified pT7T3 vector (Pharmacia).
                     Library went through one round of normalization to a Cot =
                     20. Library constructed by Bento Soares and M.Fatima
                     Bonaldo."
BASE COUNT       93 a    107 c    100 g    109 t      5 others
ORIGIN


  Query Match             8.0%;  Score 238;  DB 141;  Length 414;
  Best Local Similarity  99.7%;  Pred. No. 1.8e-114;
```

```
  Matches  288;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     306 aactgtggagaaggcgttcagagactcatgcaggagcacaagttaaaggttgctcgcctg 365
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     289 AACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTAAAGGTTGCTCGCCTG 230

Qy     366 gacaacatattcctgacacgaatgcactggtctaatgttgggggcttaagtggaatgatt 425
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     229 GACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGCTTAAGTGGAATGATT 170

Qy     426 cttactttaaaggaaaccgggcttccaaagtgtgtactttctggacctccacaactggaa 485
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     169 CTTACTTTAAAGGAAACCGGGCTTCNAAAGTGTGTACTTTCTGGACCTCCACAACTGGAA 110

Qy     486 aaatacctcgaagcaatcaaaatattttctggtccattgaaaggaatagaactggctgtg 545
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     109 AAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGAATAGAACTGGCTGTG 50

Qy     546 cggccccactctgccccagaatacgaggatgaaaccatgacagtttacc 594
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49 CGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTTTACC 1


RESULT  74
BE409312
LOCUS       BE409312      691 bp    mRNA            EST       21-JUL-2000
DEFINITION  601300940F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3635576 5',
            mRNA sequence.
ACCESSION   BE409312
VERSION     BE409312.1  GI:9345762
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 691)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM331  row: p  column: 09
            High quality sequence start: 4
            High quality sequence stop: 615.
FEATURES             Location/Qualifiers
     source          1. .691
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3635576"
                     /clone_lib="NIH_MGC_21"
                     /tissue_type="choriocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Size-selected >500bp
                     for average insert size 1.8kb.  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      164 a    198 c    197 g    132 t
ORIGIN


  Query Match             8.0%;  Score 236;  DB 107;  Length 691;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-113;
  Matches  236;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      32 ttttggtggagacgggcgcatgtgggcgctttgctcgctgctgcggtccgcggccggacg 91
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      24 TTTTGGTGGAGACGGGCGCATGTGGGCGCTTTGCTCGCTGCTGCGGTCCGCGGCCGGACG 83

Qy      92 caccatgtcgcagggacgcaccatatcgcaggcacccgcccgccgcgagcggccgcgcaa 151
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      84 CACCATGTCGCAGGGACGCACCATATCGCAGGCACCCGCCCGCCGCGAGCGGCCGCGCAA 143

Qy     152 ggacccgctgcggcacctgcgcacgcgagagaagcgcggaccgtcggggtgctccggcgg 211
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     144 GGACCCGCTGCGGCACCTGCGCACGCGAGAGAAGCGCGGACCGTCGGGGTGCTCCGGCGG 203

Qy     212 cccaaacaccgtgtacctgcaggtggtggcagcgggtagccgggactcgggcgccg 267
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     204 CCCAAACACCGTGTACCTGCAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCG 259


RESULT  75
AI201492/c
LOCUS       AI201492      439 bp    mRNA            EST       30-OCT-1998
DEFINITION  qs74b03.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:1943789 3',
            mRNA sequence.
ACCESSION   AI201492
VERSION     AI201492.1  GI:3754098
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 439)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            unknown library type
            Insert Length: 717   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 407.
FEATURES             Location/Qualifiers
     source          1. .439
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1943789"
                     /clone_lib="NCI_CGAP_Pr28"
                     /sex="male"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Plasmid DNA from the
                     normalized library NCI_CGAP_Pr22 was prepared, and ss
                     circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     985608-986759, 1101192-1101959, and 1217928-1220615).
                     Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT       93 a    119 c    109 g    118 t
ORIGIN


  Query Match             7.9%;  Score 235;  DB 17;  Length 439;
  Best Local Similarity   99.1%;  Pred. No. 7.1e-113;
  Matches  435;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy    2518 tcagagcccagtgaagatctgggagaccctgaactcagaaggctgtgtgtcttctgcccc 2577
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     439 TCAGAGCCCAGTGAAGATCTGGGAGACCCTGAACTCAGAAGGCTGTGTGTCTTCTGCCCC 380
```

```
Qy   2578 acgcacgcacccgtatctgccctccttgctggtagaagctgaagagcacggtcccccagg 2637
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    379 ACGCACGCACCCGTATCTGCCCTCCTTGCTGGTAGAAGCTGAAGAGCACGGTCCCCCAGG 320

Qy   2638 aggcagctcaggataggtggtatggagctgtgccgaggcttgggctcccacataagcact 2697
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    319 AGGCAGCTCAGGATAGGTGGTATGGAGCTGTGCCGAGGCTTGGGGTCCCACATAAGCACT 260

Qy   2698 agtctatagatgcctcttaggactggtgcctggcacagccgcgggccaggaggctgccac 2757
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    259 AGTCTATAGATGCCTCTTAGGACTGGTGCCTGGCACAGCTGCGGGCCAGGAGGCTGCCAC 200

Qy   2758 acggaagcaagcagatgaactaatttcatttcaaggcagtttttaaagaagtcttggaaa 2817
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    199 ACGGAAGCAAGCAGATGAACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCATGGAAA 140

Qy   2818 cagacggcggcacctttcctctaatccagcaaagtgattccctgcacaccagagacaagc 2877
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    139 CAGACGGCGGCACCTTTCCTCTAATCCAGCAAAATGATTCCCTGCACACCAGAGACAAGC 80

Qy   2878 agagtaacaggatcagtgggtctaagtgtccgagacttaacgaaaatagtatttcagctg 2937
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     79 AGAGTAACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTG 20

Qy   2938 caataaagattgagtttgc 2956
          ||||||||||||||||||
Db     19 CAATAAAGATTGAGTTTGC 1
```

RESULT 76
AF188525/c
LOCUS AF188525 233 bp mRNA EST 02-MAR-2000
DEFINITION AF188525 Homo sapiens ATCC HTB-12; SW1088 Homo sapiens cDNA clone IRIG 1, mRNA sequence.
ACCESSION AF188525
VERSION AF188525.1 GI:7144571
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 233)
AUTHORS Ye,Z. and Connor,J.R.
TITLE Identification of Iron Regulated Genes by Rescreening cDNA Libraries from SSH with Antisense Probe from Three Iron Conditions
JOURNAL Unpublished (2000)
COMMENT Contact: Ye Z
Neuroscience and Anatomy
Pennsylvania State University College of Medicine
500 University Drive, Hershey, PA 17033, USA
library screened by SSH and reverse Northern blot; decreased expression in iron loading was confirmed by Northern blot.
FEATURES Location/Qualifiers
source 1..233
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IRIG 1"
/clone_lib="Homo sapiens ATCC HTB-12; SW1088"
/tissue_type="astrocytoma"
/cell_line="ATCC HTB-12; SW1088"
BASE COUNT 56 a 68 c 55 g 54 t
ORIGIN

Query Match 7.9%; Score 233; DB 14; Length 233;
Best Local Similarity 100.0%; Pred. No. 7.8e-112;
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    226 acctgcaggtggtggcagcgggtagccgggactcgggcgccgcgctctacgtcttctccg 285
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    233 ACCTGCAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCG 174

Qy    286 agttcaaccggtatctcttcaactgtggagaaggcgttcagagactcatgcaggagcaca 345
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    173 AGTTCAACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACA 114

Qy    346 agttaaaggttgctcgcctggacaacatattcctgacacgaatgcactggtctaatgttg 405
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    113 AGTTAAAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTG 54

Qy    406 ggggcttaagtggaatgattcttactttaaaggaaaccgggcttccaaagtgt 458
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     53 GGGGCTTAAGTGGAATGATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGT 1
```

RESULT 77
H14462/c
LOCUS H14462 394 bp mRNA EST 27-JUN-1995
DEFINITION yl25f04.r1 Soares breast 3NbHBst Homo sapiens cDNA clone IMAGE:159295 5', mRNA sequence.
ACCESSION H14462
VERSION H14462.1 GI:879282
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 394)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 509
High quality sequence stops: 268
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 509 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 268.
FEATURES Location/Qualifiers
source 1..394
/organism="Homo sapiens"
/db_xref="GDB:578182"
/db_xref="taxon:9606"
/clone="IMAGE:159295"
/clone_lib="Soares breast 3NbHBst"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: breast; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 20. Library constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 90 a 96 c 99 g 107 t 2 others
ORIGIN

Query Match 7.9%; Score 233; DB 141; Length 394;
Best Local Similarity 100.0%; Pred. No. 8.1e-112;
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    362 cctggacaacatattcctgacacgaatgcactggtctaatgttgggggcttaagtggaat 421
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    233 CCTGGACAACATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGCTTAAGTGGAAT 174

Qy    422 gattcttactttaaaggaaaccgggcttccaaagtgtgtactttctggacctccacaact 481
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    173 GATTCTTACTTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGACCTCCACAACT 114

Qy    482 ggaaaaatacctcgaagcaatcaaaatattttctggtccattgaaaggaatagaactggc 541
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    113 GGAAAAATACCTCGAAGCAATCAAAATATTTTCTGGTCCATTGAAAGGAATAGAACTGGC 54

Qy    542 tgtgcggccccactctgccccagaatacgaggatgaaaccatgacagtttacc 594
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     53 TGTGCGGCCCCACTCTGCCCCAGAATACGAGGATGAAACCATGACAGTTTACC 1


RESULT  78
AA522537/c
LOCUS       AA522537      865 bp    mRNA            EST       20-AUG-1997
DEFINITION  ni38e08.s1 NCI_CGAP_Lu1 Homo sapiens cDNA clone IMAGE:979142 3',
            mRNA sequence.
ACCESSION   AA522537
VERSION     AA522537.1  GI:2263249
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 865)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Stratagene, Inc., David B. Krizman,
            Ph.D.
             cDNA Library Arraying: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 436.
FEATURES             Location/Qualifiers
     source          1. .865
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:979142"
                     /clone_lib="NCI_CGAP_Lu1"
                     /tissue_type="lung tumor"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: lung; Vector: Bluescript SK-; Site_1: EcoRI;
                     Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo dT.
                     Bulk lung tumor. 5' adaptor sequence: 5' GAATTCGGCACGAG 3'
                     3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'
                     Average insert size:  1.1 kb."
BASE COUNT      179 a    235 c    227 g    218 t      6 others
ORIGIN


  Query Match           7.8%;  Score 232;  DB 8;  Length 865;
  Best Local Similarity   99.5%;  Pred. No. 2.9e-111;
  Matches  402;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy   2540 gagaccctgaactcagaaggctgtgtgtctttctgccccacgcacgcacccgtatctgccc 2599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    403 GAGACCCTGAACTCAGAAGGCTGTGTGTCTTTCTGCCCCACGCACGCACCCGTATCTGCCC 344

Qy   2600 tccttgctggtagaagctgaagagcacggtcccccaggaggcagctcaggataggtggta 2659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    343 TCCTTGCTGGTAGAAGCTGAAGAGCACGGTCCCCCAGGAGGCAGCTCAGGATAGGTGGTA 284

Qy   2660 tggagctgtgccgaggcttgggctcccacataagcactagtctatagatgcctcttagga 2719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    283 TGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTCTTAGGA 224

Qy   2720 ctggtgcctggcacagccgcgggccaggaggctgccacacggaagcaagcagatgaacta 2779
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    223 CTGGTGCCT-GCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGAACTA 165

Qy   2780 atttcatttcaaggcagtttttaaagaagtcttggaaacagacggcggcacctttcctct 2839
          |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db    164 ATTTCATTTCAAGGCAGTTTTTAAAGAAGTCATGGAAACAGACGGCGGCACCTTTCCTCT 105

Qy   2840 aatccagcaaagtgattccctgcacaccagagacaagcagagtaacaggatcagtgggtc 2899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    104 AATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTGGGTC 45

Qy   2900 taagtgtccgagacttaacgaaaatagtatttcagctgcaataa 2943
          |||||||||||||||||||||||||||||||||||||||||||
Db     44 TAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAA 1


RESULT  79
AA928608/c
LOCUS       AA928608      282 bp    mRNA            EST       07-JUL-1998
DEFINITION  om75b03.s1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1552973 3',
            mRNA sequence.
ACCESSION   AA928608
VERSION     AA928608.1  GI:3076899
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 282)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 846    Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 255.
FEATURES             Location/Qualifiers
     source          1. .282
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1552973"
                     /clone_lib="NCI_CGAP_GC4"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from 3 pooled
                     germ cell tumors, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  Library is normalized. Library was constructed by
                     Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       63 a     68 c     66 g     84 t      1 others
```

```
ORIGIN

  Query Match             7.7%;  Score 228;  DB 13;  Length 282;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-109;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2730 gcacagccgcgggccaggaggctgccacacggaagcaagcagatgaactaatttcatttc 2789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    228 GCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGAACTAATTTCATTTC 169

Qy   2790 aaggcagtttttaaagaagtcttggaaacagacggcggcacctttcctctaatccagcaa 2849
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    168 AAGGCAGTTTTTAAAGAAGTCTTGGAAACAGACGGCGGCACCTTTCCTCTAATCCAGCAA 109

Qy   2850 agtgattccctgcacaccagagacaagcagagtaacaggatcagtgggtctaagtgtccg 2909
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    108 AGTGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTGGGTCTAAGTGTCCG 49

Qy   2910 agacttaacgaaaatagtatttcagctgcaataaagattgagtttgca 2957
          |||||||||||||||||||||||||||||||||||||||||||||||
Db     48 AGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTGCA 1


RESULT  80
T34024
LOCUS       T34024         282 bp    mRNA            EST       06-SEP-1995
DEFINITION  EST61387 Human White blood cells Homo sapiens cDNA 5' end similar
            to None, mRNA sequence.
ACCESSION   T34024
VERSION     T34024.1  GI:616122
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 282)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
            ,C.J., Lee,N., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
            ,O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W., Clayton,R.A.,
            Cline,R.T., Cotton,M.D., Earle-Hughes,J., Fine,L.D., FitzGerald
            ,L.M., FitzHugh,W.M., Fritchman,J.L., Geoghagen,N.S.M., Glodek,A.,
            Gnehm,C.L., Hanna,M.C., Hedblom,E., HinkleJr,P.S., Kelley,J.M.,
            Klimek,K.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pellegrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.-J.,
            Dimke,D., Feng,P., Ferrie,A., Fischer,C., Hastings,G.A., He,W.-W.,
            Hu,J.-S., Greene,J.M., Gruber,J., Hudson,P., Kim,A., Kozak,D.L.,
            Kunsch,C., Ji,H., Li,H., Meissner,P.S., Olsen,H., Raymond,L., Wei
            ,Y.-F., Wing,J., Xu,C., Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon
            ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
            Venter,J.C.
  TITLE     Initial Assessment of Human Gene Diversity and Expression Patterns
            Based Upon 83 Million Basepairs of cDNA Sequence
  JOURNAL   Nature 377, 3-174 (1995)
  MEDLINE   96026280
COMMENT     Other_ESTs: THC15444
            Contact: Venter, JC
            The Institute for Genomic Research
            932 Clopper Rd, Gaithersburg, MD 20878
            Tel: 3018699056
            Fax: 3018699423
            Email: tdbinfo@tdb.tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please contact the TIGR Database
            (tdbinfo@tdb.tigr.org)
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .282
                     /organism="Homo sapiens"
                     /db_xref="ATCC (inhost):104694"
                     /db_xref="taxon:9606"
                     /clone_lib="Human White blood cells"
                     /tissue_type="white blood cells"
                     /note="Organ: blood"
BASE COUNT       62 a     83 c     73 g     62 t      2 others
ORIGIN


  Query Match             7.7%;  Score 228;  DB 145;  Length 282;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-109;
  Matches  228;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1672 ctgtgtttgtgtcccacctgcacgcagatcaccacacgggcttgccaagtatcttgctgc 1731
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     26 CTGTGTTTGTGTCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTGC 85

Qy   1732 agagagaacgcgccttggcatctttgggaaagccgcttcacccttgctggtggttgccc 1791
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     86 AGAGAGAACGCGCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCCC 145

Qy   1792 ccaaccagctcaaagcctggctccagcagtaccacaaccagtgccaggaggtcctgcacc 1851
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    146 CCAACCAGCTCAAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCACC 205

Qy   1852 acatcagtatgattcctgccaaatgccttcaggaaggggctgagatct 1899
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    206 ACATCAGTATGATTCCTGCCAAATGCCTTCAGGAAGGGGCTGAGATCT 253


RESULT  81
AI357786/c
LOCUS       AI357786       433 bp    mRNA            EST       15-FEB-1999
DEFINITION  qu98d07.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:1980109 3',
            mRNA sequence.
ACCESSION   AI357786
VERSION     AI357786.1  GI:4109407
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 433)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 2467   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 397.
FEATURES             Location/Qualifiers
     source          1. .433
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1980109"
                     /clone_lib="NCI_CGAP_Gas4"
                     /tissue_type="poorly differentiated adenocarcinoma with
                     signet ring cell features"
                     /lab_host="DH10B"
                     /note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.69 kb. Life Technologies catalog #:
                     11549-011"
```

```
ASE COUNT       92 a    117 c    105 g    119 t
RIGIN

  Query Match             7.7%;  Score 227;  DB 19;  Length 433;
  Best Local Similarity   99.1%;  Pred. No. 1.2e-108;
  Matches  427;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

y   2528 gtgaagatctgggagaccctgaactcagaaggctgtgtgtcttctgccccacgcacgcac 2587
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    433 GTGAAGATCTGGGAGACCCTGAACTCAGAAGGCTGTGTGTCTTCTGCCCCACGCACGCAC 374

y   2588 ccgtatctgccctccttgctggtagaagctgaagagcacggtccccccaggaggcagctca 2647
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    373 CCGTATCTGCCCTCCTTGCTGGTAGAAGCTGAAGAGCACGGTCCCCCAGGAGGCAGCTCA 314

y   2648 ggataggtggtatggagctgtgccgaggcttgggctcccacataagcactagtctataga 2707
       |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
b    313 GGATAGGTGGTATGGAGCTGTGCCGAGGCTTGGGGTCCCACATAAGCACTAGTCTATAGA 254

y   2708 tgcctcttaggactggtgcctggcacagccgcgggccaggaggctgccacacggaagcaa 2767
       |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
b    253 TGCCTCTTAGGACTGGTGCCTGGCACAGCTGCGGGCCAGGAGGCTGCCACACGGAAGCAA 194

y   2768 gcagatgaactaatttcatttcaaggcagtttttaaagaagtcttggaaacagacggcgg 2827
       ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
b    193 GCAGATGAACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCATGGAAACAGACGGCGG 134

y   2828 caccttttcctctaatccagcaaagtgattccctgcacaccagagacaagcagagtaacag 2887
       |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
b    133 CACCTTTCCTCTAATCCAGCAAAATGATTCCCTGCACACCAGAGACAAGCAGAGTAACAG 74

y   2888 gatcagtgggtctaagtgtccgagacttaacgaaaatagtatttcagctgcaataaagat 2947
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b     73 GATCAGTGGGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGAT 14

y   2948 tgagtttgcaa 2958
       ||||||||||||
b     13 TGAGTTTGCAA 3


ESULT  82
E615669
OCUS       BE615669      872 bp    mRNA            EST       24-AUG-2000
EFINITION  601279347F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3611338 5',
           mRNA sequence.
CCESSION   BE615669
ERSION     BE615669.1  GI:9897268
EYWORDS    EST.
OURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
EFERENCE   1  (bases 1 to 872)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
OMMENT     Contact: Robert Strausberg, Ph.D.
           Tel: (301) 496-1550
           Email: Robert_Strausberg@nih.gov
           Tissue Procurement: ATCC
           cDNA Library Preparation: Ling Hong/Rubin Laboratory
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLCM268  row: n  column: 11
           High quality sequence stop: 667.
EATURES             Location/Qualifiers
     source         1. .872
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:3611338"
                    /clone_lib="NIH_MGC_39"
                    /tissue_type="adenocarcinoma"
                    /lab_host="DH10B (phage-resistant)"
                    /note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
                    Site_2: EcoRI; cDNA made by oligo-dT priming.
                    Directionally cloned into EcoRI/XhoI sites using the
                    following 5' adaptor: GGCACGAG(G).  Library constructed
                    by Ling Hong in the laboratory of Gerald M. Rubin
                    (University of California,  Berkeley) using ZAP-cDNA
                    synthesis kit (Stratagene) and Superscript II RT (Life
                    Technologies)."
BASE COUNT      196 a    239 c    257 g    180 t
ORIGIN

  Query Match             7.6%;  Score 224;  DB 110;  Length 872;
  Best Local Similarity   100.0%;  Pred. No. 5e-107;
  Matches  224;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    213 ccaaacaccgtgtacctgcaggtggtggcagcgggtagccgggactcgggcgccgcgctc 272
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    208 CCAAACACCGTGTACCTGCAGGTGGTGGCAGCGGGTAGCCGGGACTCGGGCGCCGCGCTC 267

Qy    273 tacgtcttctccgagttcaaccggtatctcttcaactgtggagaaggcgttcagagactc 332
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    268 TACGTCTTCTCCGAGTTCAACCGGTATCTCTTCAACTGTGGAGAAGGCGTTCAGAGACTC 327

Qy    333 atgcaggagcacaagttaaaggttgctcgcctggacaacatattcctgacacgaatgcac 392
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    328 ATGCAGGAGCACAAGTTAAAGGTTGCTCGCCTGGACAACATATTCCTGACACGAATGCAC 387

Qy    393 tggtctaatgttgggggcttaagtggaatgattcttactttaaa 436
          |||||||||||||||||||||||||||||||||||||||||||
Db    388 TGGTCTAATGTTGGGGGCTTAAGTGGAATGATTCTTACTTTAAA 431


RESULT  83
AA235532
LOCUS       AA235532      291 bp    mRNA            EST       08-AUG-1997
DEFINITION  zt29g11.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone
            IMAGE:723812 3' similar to SW:YK59_YEAST P36159 HYPOTHETICAL 96.8
            KD PROTEIN IN SIS2-MTD1 INTERGENIC REGION. ;, mRNA sequence.
ACCESSION   AA235532
VERSION     AA235532.1  GI:1860004
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 291)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: similarity on wrong strand
            Possible reversed clone: polyT not found
            Insert Length: 1829   Std Error: 0.00
            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 271.
```

```
FEATURES             Location/Qualifiers
     source          1. .291
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:723812"
                     /clone_lib="Soares ovary tumor NbHOT"
                     /sex="Female"
                     /tissue_type="ovarian tumor"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: ovary; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo. "
BASE COUNT       64 a     73 c     88 g     66 t
ORIGIN


  Query Match             7.5%;  Score 221;  DB 4;  Length 291;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-105;
  Matches  221;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    251 ccgggactcgggcgccgcgctctacgtcttctccgagttcaaccggtatctcttcaactg 310
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     71 CCGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTCAACCGGTATCTCTTCAACTG 130

Qy    311 tggagaaggcgttcagagactcatgcaggagcacaagttaaaggttgctcgcctggacaa 370
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    131 TGGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTAAAGGTTGCTCGCCTGGACAA 190

Qy    371 catattcctgacacgaatgcactggtctaatgttgggggcttaagtggaatgattcttac 430
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    191 CATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGCTTAAGTGGAATGATTCTTAC 250

Qy    431 tttaaaggaaaccgggcttccaaagtgtgtactttctggac 471
          |||||||||||||||||||||||||||||||||||||||||
Db    251 TTTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGGAC 291


RESULT  84
AI804749/c
LOCUS       AI804749      477 bp    mRNA            EST       07-MAR-2000
DEFINITION  tu42d02.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:2253699 3',
            mRNA sequence.
ACCESSION   AI804749
VERSION     AI804749.1  GI:5370221
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 477)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 876  Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 458.
FEATURES             Location/Qualifiers
     source          1. .477
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2253699"
                     /clone_lib="NCI_CGAP_Pr28"
                     /sex="male"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Plasmid DNA from the
                     normalized library NCI_CGAP_Pr22 was prepared, and ss
                     circles were made in vitro. Following HAP purification,
                     this DNA was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from a pool
                     of 5,000 clones made from the same library (cloneIDs
                     985608-986759, 1101192-1101959, and 1217928-1220615).
                     Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       93 a    130 c    120 g    133 t      1 others
ORIGIN


  Query Match             7.4%;  Score 220;  DB 25;  Length 477;
  Best Local Similarity   98.9%;  Pred. No. 6.3e-105;
  Matches  470;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy   2484 cgggcccacacagaggagccacaggccaagaaggtcagagcccagtgaagatctgggaga 2543
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    477 CGGGCCCACACAGAGGAGCCACAGGCCAAGAAGGTCAGAGCCCAGTGAAGATCTGGGAGA 418

Qy   2544 ccctgaactcagaaggctgtgtgtcttctgccccacgcacgcacccgtatctgccctcct 2603
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    417 CCCTGAACTCAGAAGGCTGTGTGTCTTCTGCCCCACGCACGCACCCGTATCTGCCCTCCT 358

Qy   2604 tgctggtagaagctgaagagcacggtcccccaggaggcagctcaggataggtggtatgga 2663
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    357 TGCTGGTAGAAGCTGAAGAGCACGGTCCCCCAGGAGGCAGCTCAGGATAGGTGGTATGGA 298

Qy   2664 gctgtgccgaggcttgggctcccacataagcactagtctatagatgcctcttaggactgg 2723
          |||||||| |||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    297 GCTGTGCCNAGGCTTGGGGTCCCACATAAGCACTAGTCTATAGATGCCTCTTAGGACTGG 238

Qy   2724 tgcctggcacagccgcgggccaggaggctgccacacggaagcaagcagatgaactaattt 2783
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    237 TGCCTGGCACAGCTGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGAACTAATTT 178

Qy   2784 catttcaaggcagtttttaaagaagtcttggaaacagacggcggcacctttcctctaatc 2843
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    177 CATTTCAAGGCAGTTTTTAAAGAAGTCATGGAAACAGACGGCGGCACCTTTCCTCTAATC 118

Qy   2844 cagcaaagtgattccctgcacaccagagacaagcagagtaacaggatcagtgggtctaag 2903
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    117 CAGCAAAATGATTCCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTGGGTCTAAG 58

Qy   2904 tgtccgagacttaacgaaaatagtatttcagctgcaataaagattgagtttgcaa 2958
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     57 TGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTGCAA 3


RESULT  85
AW296524
LOCUS       AW296524      228 bp    mRNA            EST       16-JAN-2000
DEFINITION  UI-H-BW0-aiv-a-09-0-UI.s1 NCI_CGAP_Sub6 Homo sapiens cDNA clone
            IMAGE:2730521 3', mRNA sequence.
ACCESSION   AW296524
VERSION     AW296524.1  GI:6703160
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 228)
```

```
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message. cDNA Library Preparation: M.B.
            Soares Lab Clone distribution: NCI-CGAP clone distribution
            information can be found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: M13 Forward
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .228
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2730521"
                     /clone_lib="NCI_CGAP_Sub6"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; NCI_CGAP_Sub6
                     is a subtracted library derived from BW, which consists of
                     a  mixture of four normalized libraries: NCI_CGAP_Brn50,
                     NCI_CGAP_Lu13,  NCI_CGAP_Ov18, GBC1 . The NCI_CGAP_Sub6
                     library had 7 million recombinants.  A single-stranded DNA
                     preparation of BW was used as a tracer in a subtractive
                     hybridization with a driver comprising: the IMAGE pool
                     (NCI_CGAP_Kid3 pool 1  LLAM 3334-3337, 3682-3683,
                     3798-3803 (IMAGE CloneIDs 1322376-1323911,
                     1456008-1456775,1500552-1502855); NCI_CGAP_Kid5 pool 1
                     LLAM 3338-3342,3722-3725, 3776-3778 (IMAGE CloneIDs
                     1323912-1325831, 1471368-1472903, 1492104-1493255);
                     NCI_CGAP_Lu5 pool 1 LLAM 3575-3582, 3851-3854 (IMAGE
                     CloneIDs 1414920-1417991, 1520904-1522439); NCI_CGAP_GC4
                     pool 1 LLAM 3164-3167, 3716-3720, 3733-3735  (IMAGE
                     CloneIDs 1257096-1258631,1469064-1470983, 1475592-1476743
                     ); NCI_CGAP_Pr22 pool 1 LLAM 2457-2459, 2758-2759,
                     3062-3068 (IMAGE CloneIDs 985608-986759, 1101192-1101959,
                     1217928-1220615); NCI_CGAP_Co10 pool 1 LLAM 2644-2653,
                     2871-2872 (IMAGE CloneIDs 1057416-1061255,1144584-1145351
                     ). (50% of the driver  population), plus a pool of 3,840
                     arrayed clones from  NCI_CGAP_Sub1 (IMAGE  CloneIDs
                     2708616-2710535) and NCI_CGAP_Sub2 (IMAGE CloneIDs
                     2710536-2712455)    (20% of the driver population), plus a
                     pool of 11,136 clones from NCI_CGAP_Sub3  (IMAGE CloneIDs
                     2712456-2723591) (30% of the driver population).
                     Subtraction was performed as previously described [Bonaldo
                     , Lennon & Soares (1996): Normalization and  Subtraction:
                     Two Approaches To Facilitate Gene Discovery. Genome
                     Research 6, 791-806.
                     TAG_LIB=NCI_CGAP-Lu13
                     TAG_TISSUE=lung
                     TAG_SEQ=GCCGG"
BASE COUNT       41 a     72 c     78 g     37 t
ORIGIN


  Query Match             7.4%;  Score 218;  DB 88;  Length 228;
  Best Local Similarity   100.0%;  Pred. No. 6.9e-104;
  Matches  218;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    141 cggccgcgcaaggacccgctgcggcacctgcgcacgcgagagaagcgcggaccgtcgggg 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CGGCCGCGCAAGGACCCGCTGCGGCACCTGCGCACGCGAGAGAAGCGCGGACCGTCGGGG 60

Qy    201 tgctccggcggcccaaacaccgtgtacctgcaggtggtggcagcgggtagccgggactcg 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGCTCCGGCGGCCCAAACACCGTGTACCTGCAGGTGGTGGCAGCGGGTAGCCGGGACTCG 120

Qy    261 ggcgccgcgctctacgtcttctccgagttcaaccggtatctcttcaactgtggagaaggc 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    121 GGCGCCGCGCTCTACGTCTTCTCCGAGTTCAACCGGTATCTCTTCAACTGTGGAGAAGGC 180

Qy    321 gttcagagactcatgcaggagcacaagttaaaggttgc 358
          ||||||||||||||||||||||||||||||||||||||
Db    181 GTTCAGAGACTCATGCAGGAGCACAAGTTAAAGGTTGC 218


RESULT  86
N36229
LOCUS       N36229         448 bp    mRNA            EST       16-JAN-1996
DEFINITION  yy30c04.s1 Soares melanocyte 2NbHM Homo sapiens cDNA clone
            IMAGE:272742 3' similar to SW:YK59_YEAST P36159 HYPOTHETICAL 96.8
            KD PROTEIN IN SIS2-MTD1 INTERGENIC REGION. [1] ;, mRNA sequence.
ACCESSION   N36229
VERSION     N36229.1  GI:1157371
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 448)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 365
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: similarity on wrong strand
            Possible reversed clone: polyT not found
            Seq primer: m13 -40 forward
            High quality sequence stop: 365.
FEATURES             Location/Qualifiers
     source          1. .448
                     /organism="Homo sapiens"
                     /db_xref="GDB:3882384"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:272742"
                     /clone_lib="Soares melanocyte 2NbHM"
                     /sex="Male"
                     /tissue_type="melanocyte"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCAGTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library constructed by Bento Soares and
                     M.Fatima Bonaldo. RNA from normal foreskin melanocytes
                     (FS374) was kindly provided by Dr. Anthony P. Albino."
BASE COUNT      110 a    113 c    127 g     94 t      4 others
ORIGIN


  Query Match             7.4%;  Score 218;  DB 142;  Length 448;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-104;
  Matches  218;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    252 cgggactcgggcgccgcgctctacgtcttctccgagttcaaccggtatctcttcaactgt 311
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 CGGGACTCGGGCGCCGCGCTCTACGTCTTCTCCGAGTTCAACCGGTATCTCTTCAACTGT 168
```

```
Qy    312 ggagaaggcgttcagagactcatgcaggagcacaagttaaaggttgctcgcctggacaac 371
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 GGAGAAGGCGTTCAGAGACTCATGCAGGAGCACAAGTTAAAGGTTGCTCGCCTGGACAAC 228

Qy    372 atattcctgacacgaatgcactggtctaatgttgggggcttaagtggaatgattcttact 431
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 ATATTCCTGACACGAATGCACTGGTCTAATGTTGGGGGCTTAAGTGGAATGATTCTTACT 288

Qy    432 ttaaaggaaaccgggcttccaaagtgtgtactttctgg 469
          |||||||||||||||||||||||||||||||||||
Db    289 TTAAAGGAAACCGGGCTTCCAAAGTGTGTACTTTCTGG 326


RESULT  87
AW511765/c
LOCUS       AW511765      316 bp    mRNA            EST       03-MAR-2000
DEFINITION  xu76f03.x1 NCI_CGAP_Kid8 Homo sapiens cDNA clone IMAGE:2807645 3',
            mRNA sequence.
ACCESSION   AW511765
VERSION     AW511765.1  GI:7149843
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 316)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            image.llnl.gov/image/html/iresources.shtml
            Seq primer: -40UP from Gibco
            High quality sequence stop: 239.
FEATURES             Location/Qualifiers
     source          1. .316
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2807645"
                     /clone_lib="NCI_CGAP_Kid8"
                     /tissue_type="renal cell tumor"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.2 kb. Life Technologies catalog #:
                     11524-014"
BASE COUNT       69 a     84 c     72 g     90 t      1 others
ORIGIN


  Query Match             7.2%;  Score 214;  DB 91;  Length 316;
  Best Local Similarity   99.6%;  Pred. No. 9.3e-102;
  Matches  264;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   2692 agcactagtctatagatgcctcttaggactggtgcctggcacagccgcgggccaggaggc 2751
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    265 AGCACTAGTCTATAGATGCCTCTTAGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGC 206

Qy   2752 tgccacacggaagcaagcagatgaactaatttcatttcaaggcagtttttaaagaagtct 2811
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    205 TGCCACACGGAAGCAAGCAGATGAACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCA 146

Qy   2812 tggaaacagacggcggcacctttcctctaatccagcaaagtgattccctgcacaccagag 2871
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    145 TGGAAACAGACGGCGGCACCTTTCCTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAG 86

Qy   2872 acaagcagagtaacaggatcagtgggtctaagtgtccgagacttaacgaaaatagtattt 2931
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     85 ACAAGCAGAGTAACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTT 26

Qy   2932 cagctgcaataaagattgagtttgc 2956
          |||||||||||||||||||||||||
Db     25 CAGCTGCAATAAAGATTGAGTTTGC 1


RESULT  88
AW247380
LOCUS       AW247380      249 bp    mRNA            EST       07-JAN-2000
DEFINITION  2820640.5prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2820640 5',
            mRNA sequence.
ACCESSION   AW247380
VERSION     AW247380.1  GI:6590373
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 249)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: 2820640.3prime
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling
            Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
            Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
            project Clone distribution: MGC clone distribution information can
            be found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality
            Scores: PHRED from University of Washington Genome Center. Vector
            Trimming: cross_match from University of Washington Genome Center
            PHRAP suite. Poly-T Identification: patMatch.pl from Berkeley
            Drosophila Genome Project. University of Washington Genome Center:
            http://www.genome.washington.edu
            Plate: LLCM4  row: L  column: 17
            High quality sequence stop: 225.
FEATURES             Location/Qualifiers
     source          1. .249
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2820640"
                     /clone_lib="NIH_MGC_7"
                     /tissue_type="small cell carcinoma"
                     /cell_line="MGC3"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT       39 a     83 c     94 g     33 t
ORIGIN


  Query Match             7.1%;  Score 211;  DB 88;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-100;
  Matches  211;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 cgcgggcgtaggtgaccggcggctttctcagttttggtggagacgggcgcatgtgggcgc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     20 CGCGGGCGTAGGTGACCGGCGGCTTTCTCAGTTTTGGTGGAGACGGGCGCATGTGGGCGC 79
```

```
Qy     61 tttgctcgctgctgcggtccgcggccggacgcaccatgtcgcagggacgcaccatatcgc 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     80 TTTGCTCGCTGCTGCGGTCCGCGGCCGGACGCACCATGTCGCAGGGACGCACCATATCGC 139

Qy    121 aggcacccgcccgccgcgagcggccgcgcaaggacccgctgcggcacctgcgcacgcgag 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    140 AGGCACCCGCCCGCCGCGAGCGGCCGCGCAAGGACCCGCTGCGGCACCTGCGCACGCGAG 199

Qy    181 agaagcgcggaccgtcggggtgctccggcgg 211
          |||||||||||||||||||||||||||||||
Db    200 AGAAGCGCGGACCGTCGGGGTGCTCCGGCGG 230


RESULT  89
AW378247/c
LOCUS       AW378247      650 bp    mRNA            EST       04-FEB-2000
DEFINITION  RC1-HT0217-201199-022-h08 HT0217 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW378247
VERSION     AW378247.1  GI:6882906
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 650)
  AUTHORS   HCGP http://www.ludwig.org.br/ORESTES.
  TITLE     The FAPESP/LICR Human Cancer Genome Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC1&t2=RC1-HT0217-
            201199-022-h08&t3=1999-11-20&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 74.
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HT0217"
                     /dev_stage="Adult"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT      124 a    192 c    186 g    148 t
ORIGIN


  Query Match             7.1%;  Score 209;  DB 89;  Length 650;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-99;
  Matches  209;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2291 tgaccacatgaaggtctgctttggagactttccaacaatgcccaagctgattcccccact 2350
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    579 TGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACAATGCCCAAGCTGATTCCCCCACT 520

Qy   2351 gaaagccctgtttgctggcgacatcgaggagatggaggagcgcagggagaagcgggagct 2410
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    519 GAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGGAGGAGCGCAGGGAGAAGCGGGAGCT 460

Qy   2411 gcggcaggtgcgggcggccctcctgtccagggagctggcaggcggcctggaggatgggga 2470
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    459 GCGGCAGGTGCGGGCGGCCCTCCTGTCCAGGGAGCTGGCAGGCGGCCTGGAGGATGGGGA 400

Qy   2471 gcctcagcagaagcgggcccacacagagg 2499
          ||||||||||||||||||||||||||||
Db    399 GCCTCAGCAGAAGCGGGCCCACACAGAGG 371


RESULT  90
AA504146/c
LOCUS       AA504146      219 bp    mRNA            EST       18-AUG-1997
DEFINITION  aa59e06.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:825250 3',
            mRNA sequence.
ACCESSION   AA504146
VERSION     AA504146.1  GI:2240306
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 219)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 175.
FEATURES             Location/Qualifiers
     source          1. .219
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:825250"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER).  cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTT-3'
                     ]. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT       48 a     46 c     43 g     82 t
ORIGIN


  Query Match             7.0%;  Score 207;  DB 8;  Length 219;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-98;
  Matches  207;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2752 tgccacacggaagcaagcagatgaactaatttcatttcaaggcagtttttaaagaagtct 2811
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    219 TGCCACACGGAAGCAAGCAGATGAACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCT 160

Qy   2812 tggaaacagacggcggcacctttcctctaatccagcaaagtgattccctgcacaccagag 2871
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    159 TGGAAACAGACGGCGGCACCTTTCCTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAG 100

Qy   2872 acaagcagagtaacaggatcagtgggtctaagtgtccgagacttaacgaaaatagtattt 2931
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     99 ACAAGCAGAGTAACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTT 40

Qy   2932 cagctgcaataaagattgagtttgcaa 2958
          ||||||||||||||||||||||||||
Db     39 CAGCTGCAATAAAGATTGAGTTTGCAA 13


RESULT  91
W37591/c
LOCUS       W37591        422 bp    mRNA            EST       10-OCT-1996
DEFINITION  zc10f03.r1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
            IMAGE:321917 5', mRNA sequence.
ACCESSION   W37591
VERSION     W37591.1  GI:1319196
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 422)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 678   Std Error: 0.00
            Seq primer: mob.REGA+ET.
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="Homo sapiens"
                     /db_xref="GDB:1259575"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:321917"
                     /clone_lib="Soares_parathyroid_tumor_NbHPA"
                     /tissue_type="parathyroid tumor"
                     /dev_stage="adult"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia
                     ) with a modified polylinker; Site_1: Not I; Site_2: Eco
                     RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                     primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCACCAATTTTTTTTTTTTTTTTTTT
                     TTTTT-3'], double-stranded cDNA was size selected, ligated
                     to Eco RI adapters (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of a modified pT7T3
                     vector (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid
                     adenomas was kindly provided by Dr. Stephen Marx, National
                     Institute of Diabetes and Digestive and Kidney Diseases,
                     NIH."
BASE COUNT      100 a    115 c    102 g    104 t      1 others
ORIGIN


  Query Match             7.0%;  Score 206;  DB 146;  Length 422;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-97;
  Matches  206;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    841 tcccagttgggacagctgccatcgctcccatcattgctgctgtcaaggacgggaaaagca 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    271 TCCCAGTTGGGACAGCTGCCATCGCTCCCATCATTGCTGCTGTCAAGGACGGGAAAAGCA 212

Qy    901 tcactcatgaaggaagagagattttggctgaagagctgtgtactcctccagatcctggtg 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    211 TCACTCATGAAGGAAGAGAGATTTTGGCTGAAGAGCTGTGTACTCCTCCAGATCCTGGTG 152

Qy    961 ctgcttttgtggtggtagaatgtccagatgaaagcttcattcaacccatctgtgagaatg 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    151 CTGCTTTTGTGGTGGTAGAATGTCCAGATGAAAGCTTCATTCAACCCATCTGTGAGAATG 92

Qy   1021 ccacctttcagaggtaccaaggaaag 1046
          |||||||||||||||||||||||||
Db     91 CCACCTTTCAGAGGTACCAAGGAAAG 66


RESULT  92
R51138
LOCUS       R51138        472 bp    mRNA            EST       18-MAY-1995
DEFINITION  yg71c08.r1 Soares infant brain 1NIB Homo sapiens cDNA clone
            IMAGE:38752 5' similar to SP:YK59_YEAST P36159 HYPOTHETICAL 96.8 KD
            PROTEIN IN SIS2-MTD1 INTERGENIC ;, mRNA sequence.
ACCESSION   R51138
VERSION     R51138.1  GI:813040
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 472)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 327
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: M13RP1
            High quality sequence stop: 327.
FEATURES             Location/Qualifiers
     source          1. .472
                     /organism="Homo sapiens"
                     /db_xref="GDB:411293"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:38752"
                     /clone_lib="Soares infant brain 1NIB"
                     /sex="female"
                     /dev_stage="73 days post natal"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: whole brain; Vector: Lafmid BA; Site_1: Not
                     I; Site_2: Hind III; 1st strand cDNA was primed with a Not
                     I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTT 3']
                     double-stranded cDNA was ligated to Hind III adaptors
                     (Pharmacia), digested with Not I and directionally cloned
                     into the Not I and Hind III sites of the Lafmid BA vector.
                     Library went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      105 a    131 c    125 g    102 t      9 others
ORIGIN


  Query Match             6.7%;  Score 199;  DB 144;  Length 472;
  Best Local Similarity   100.0%;  Pred. No. 8.4e-94;
```

```
  Matches  199;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1430 gcaggagtacaggaggagtgcgcaggacggcccagccccagcagagaaaagaagtcagta 1489
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     43 GCAGGAGTACAGGAGGAGTGCGCAGGACGGCCCAGCCCCAGCAGAGAAAAGAAGTCAGTA 102

Qy   1490 cccagaaatcatcttccttggaacagggtctgccatcccgatgaagattcgaaatgtcag 1549
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    103 CCCAGAAATCATCTTCCTTGGAACAGGGTCTGCCATCCCGATGAAGATTCGAAATGTCAG 162

Qy   1550 tgccacacttgtcaacataagccccgacacgtctctgctactggactgtggtgagggcac 1609
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    163 TGCCACACTTGTCAACATAAGCCCCGACACGTCTCTGCTACTGGACTGTGGTGAGGGCAC 222

Qy   1610 atttgggcagctgtgccgt 1628
          ||||||||||||||||||
Db    223 ATTTGGGCAGCTGTGCCGT 241


RESULT  93
AI141263/c
LOCUS       AI141263      416 bp    mRNA            EST       05-OCT-1998
DEFINITION  qa46h05.s1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:1689849
            3', mRNA sequence.
ACCESSION   AI141263
VERSION     AI141263.1  GI:3648720
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 416)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 711   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 233.
FEATURES             Location/Qualifiers
     source          1. .416
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1689849"
                     /clone_lib="Soares_NhHMPu_S1"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
BASE COUNT       87 a    108 c     99 g    122 t
ORIGIN


  Query Match             6.7%;  Score 198;  DB 16;  Length 416;
  Best Local Similarity   99.0%;  Pred. No. 2.8e-93;
  Matches  398;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy   2557 aggctgtgtgtcttctgccccacgcacgcacccgtatctgccctccttgctggtagaagc 2616
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    416 AGGCTGTGTGTCTTCTGCCCCACGCACGCACCCGTATCTGCCCTCCTTGCTGGTAGAAGC 357

Qy   2617 tgaagagcacggtcccccaggaggcagctcaggataggtggtatggagctgtgccgaggc 2676
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    356 TGAAGAGCACGGTCCCCCAGGAGGCAGCTCAGGATAGGTGGTATGGAGCTGTGCCGAGGC 297

Qy   2677 ttgggctcccacataagcactagtctatagatgcctcttaggactggtgcctggcacagc 2736
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    296 TTGGGGTCCCACATAAGCACTAGTCTATAGATGCCTCTTAGGACTGGTGCCTGGCACAGC 237

Qy   2737 cgcgggccaggaggctgccacacggaagcaagcagatgaactaatttcatttcaaggcag 2796
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    236 TGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGATGAACTAATTTCATTTCAAGGCAG 177

Qy   2797 tttttaaagaagtcttggaaacagacggcggcacctttcctctaatccagcaaagtgatt 2856
          |||||||||||||| ||||||||||||||||||||||||||||||||||||| |||||
Db    176 TTTTTAAAGAAGTCATGGAAACAGACGGCGGCACCTTTCCTCTAATCCAGCAAAATGATT 117

Qy   2857 ccctgcacaccagagacaagcagagtaacaggatcagtgggtctaagtgtccgagactta 2916
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    116 CCCTGCACACCAGAGACAAGCAGAGTAACAGGATCAGTGGGTCTAAGTGTCCGAGACTTA 57

Qy   2917 acgaaaatagtatttcagctgcaataaagattgagtttgcaa 2958
          ||||||||||||||||||||||||||||||||||||||||
Db     56 ACGAAAATAGTATTTCAGCTGCAATAAAGATTGAGTTTGCAA 15


RESULT  94
AA346268
LOCUS       AA346268      394 bp    mRNA            EST       21-APR-1997
DEFINITION  EST52407 Greater omentum IV Homo sapiens cDNA 5' end, mRNA
            sequence.
ACCESSION   AA346268
VERSION     AA346268.1  GI:1998525
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 394)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
            ,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
            ,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
            Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
            ,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,
            Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M.,
            Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.
            , Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
            Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L.,
            Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
            ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
            Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT     Other_ESTs: THC175624
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
```

```
          Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .394
                     /organism="Homo sapiens"
                     /db_xref="ATCC (inhost):147767"
                     /db_xref="taxon:9606"
                     /clone_lib="Greater omentum IV"
                     /dev_stage="adult"
                     /note="Organ: omentum; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI"
BASE COUNT       81 a    108 c    108 g     96 t      1 others
ORIGIN


  Query Match             6.5%;  Score 193;  DB 6;  Length 394;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-90;
  Matches  193;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1683 tcccacctgcacgcagatcaccacacgggcttgccaagtatcttgctgcagagagaacgc 1742
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     77 TCCCACCTGCACGCAGATCACCACACGGGCTTGCCAAGTATCTTGCTGCAGAGAGAACGC 136

Qy   1743 gccttggcatctttgggaaagccgcttcacccttgctggtggttgccccccaaccagctc 1802
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    137 GCCTTGGCATCTTTGGGAAAGCCGCTTCACCCTTTGCTGGTGGTTGCCCCCAACCAGCTC 196

Qy   1803 aaagcctggctccagcagtaccacaaccagtgccaggaggtcctgcaccacatcagtatg 1862
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    197 AAAGCCTGGCTCCAGCAGTACCACAACCAGTGCCAGGAGGTCCTGCACCACATCAGTATG 256

Qy   1863 attcctgccaaat 1875
          |||||||||||||
Db    257 ATTCCTGCCAAAT 269


RESULT  95
AA378232
LOCUS       AA378232      276 bp    mRNA            EST       21-APR-1997
DEFINITION  EST90926 Synovial sarcoma Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION   AA378232
VERSION     AA378232.1  GI:2030571
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 276)
  AUTHORS   Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
            ,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
            ,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,
            Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
            ,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,
            Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M.,
            Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palanques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.
            , Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
            Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L.,
            Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
            ,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
            Venter,J.C.
  TITLE     Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
  JOURNAL   Nature 377 (6547 Suppl), 3-174 (1995)
  MEDLINE   96026280
COMMENT     Other_ESTs: EST114300 THC175624
            Contact: Kerlavage, AR
            Bioinformatics
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850 USA
            Tel: 3018699056
            Fax: 3018699423
            Email: arkerlav@tigr.org
            For clone availability, additional sequence and expression
            information related to this EST, please check the TIGR Human Gene
            Index (http://www.tigr.org/tdb/hgi/hgi.html)
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1. .276
                     /organism="Homo sapiens"
                     /db_xref="ATCC (inhost):182644"
                     /db_xref="taxon:9606"
                     /clone_lib="Synovial sarcoma"
                     /sex="female"
                     /tissue_type="synovial membrane"
                     /dev_stage="adult, 20 yrs"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT       70 a     75 c     79 g     50 t      2 others
ORIGIN


  Query Match             6.2%;  Score 184;  DB 6;  Length 276;
  Best Local Similarity   99.6%;  Pred. No. 7.2e-86;
  Matches  234;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   2011 tgcacacctctggctggaaagtggtctattccggggacaccatgccctgcgaggctctgg 2070
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TGCACACCTCTGGCTGGAAAGTGGTCTATTCCGGGGACACCATGCCCTGCGAGGCTCTGG 60

Qy   2071 tccggatggggaaagatgccaccctcctgatacatgaagccaccctggaagatggtttgg 2130
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     61 TCCGGATGGGGAAAGATGCCACCCTCCTGATACATGAAGCCACCCTGGAAGATGGTTTNG 120

Qy   2131 aagaggaagcagtggaaaagacacacagcacaacgtcccaagccatcagcgtggggatgc 2190
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AAGAGGAAGCAGTGGAAAAGACACACAGCACAACGTCCCAAGCCATCAGCGTGGGGATGC 180

Qy   2191 ggatgaacgcggagttcattatgctgaaccacttcagccagcgctatgccaaggt 2245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGATGAACGCGGAGTTCATTATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGT 235


RESULT  96
Z44544
LOCUS       Z44544        290 bp    mRNA            EST       14-NOV-1994
DEFINITION  HSC23E061 normalized infant brain cDNA Homo sapiens cDNA clone
            c-23e06, mRNA sequence.
ACCESSION   Z44544
VERSION     Z44544.1  GI:573683
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 290)
  AUTHORS   Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes
            ,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F.,
            Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,
            Sebastiani-Kabaktchis,C. and Tessier,A.
  TITLE     IMAGE: molecular integration of the analysis of the human genome
            and its expression
  JOURNAL   C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
  MEDLINE   95277534
COMMENT     Contact: Genethon
            Genexpress-Genethon
            Genethon Centre de recherche sur le Genome Humain
            1,rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
            Tel: 33169472800
            Fax: 33160778698
            Email: genexpress@genethon.fr
            Single read.
            Genexpress_library_idt: C; Genexpress_sequence_idt: y1c-23e06
```

```
          Seq primer: (-21)M13_universal.
FEATURES             Location/Qualifiers
     source          1. .290
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="c-23e06"
                     /clone_lib="normalized infant brain cDNA"
                     /sex="Female"
                     /tissue_type="total brain"
                     /dev_stage="3 months old"
                     /note="Organ: brain; Vector: lafmid BA; Site_1: HindIII;
                     Site_2: NotI; sex=Female; dev_stage=3 months old;
                     isolate=muscular atrophy patient;  tissue_type=total brain
                     ;  total mRNA was oligo-(dT) primed and directionally
                     cloned 5' -> 3' into the  HindIII -> NotI sites of the
                     lafmid BA vector. Clone library from B.Soares,  Psychiatry
                     Dept. Columbia University, USA.  Normalization_method:
                     Bento Soares, P.N.A.S in press"
BASE COUNT       65 a     72 c     80 g     70 t      3 others
ORIGIN


  Query Match             6.2%;  Score 182;  DB 147;  Length 290;
  Best Local Similarity   100.0%;  Pred. No. 8.3e-85;
  Matches  182;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    1026 tttcagaggtaccaaggaaaggcagatgcccccgtggccttggtggttcacatggcccca 1085
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     109 TTTCAGAGGTACCAAGGAAAGGCAGATGCCCCCGTGGCCTTGGTGGTTCACATGGCCCCA 168

Qy    1086 gcatctgtgcttgtggacagcaggtaccagcagtggatggagaggtttgggcctgacacc 1145
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     169 GCATCTGTGCTTGTGGACAGCAGGTACCAGCAGTGGATGGAGAGGTTTGGGCCTGACACC 228

Qy    1146 cagcacttggtcctgaatgagaactgtgcctcagttcacaaccttcgcagccacaagatt 1205
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     229 CAGCACTTGGTCCTGAATGAGAACTGTGCCTCAGTTCACAACCTTCGCAGCCACAAGATT 288

Qy    1206 ca 1207
           ||
Db     289 CA 290


RESULT  97
H03318/c
LOCUS        H03318         424 bp    mRNA            EST       20-JUN-1995
DEFINITION   yj47e10.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
             IMAGE:151914 3', mRNA sequence.
ACCESSION    H03318
VERSION      H03318.1  GI:866251
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 424)
  AUTHORS    Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
             ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
             Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
             ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE      The WashU-Merck EST Project
  JOURNAL    Unpublished (1995)
COMMENT      Contact: Wilson RK
             Washington University School of Medicine
             4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
             Tel: 314 286 1800
             Fax: 314 286 1810
             Email: est@watson.wustl.edu
             Insert Size: 1152
             High qality sequence stops: 315
             Source: IMAGE Consortium, LLNL
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             Insert Length: 1152   Std Error: 0.00
             Seq primer: Promega -21m13
             High quality sequence stop: 315.
FEATURES             Location/Qualifiers
     source          1. .424
                     /organism="Homo sapiens"
                     /db_xref="GDB:564064"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:151914"
                     /clone_lib="Soares placenta Nb2HP"
                     /sex="Female"
                     /dev_stage="placenta obtained at birth (full term)"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo.  "
BASE COUNT       89 a    105 c    116 g    106 t      8 others
ORIGIN


  Query Match             6.2%;  Score 182;  DB 141;  Length 424;
  Best Local Similarity   100.0%;  Pred. No. 8.5e-85;
  Matches  182;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    2762 aagcaagcagatgaactaatttcatttcaaggcagtttttaaagaagtcttggaaacaga 2821
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     195 AAGCAAGCAGATGAACTAATTTCATTTCAAGGCAGTTTTTAAAGAAGTCTTGGAAACAGA 136

Qy    2822 cggcggcacctttcctctaatccagcaaagtgattccctgcacaccagagacaagcagag 2881
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135 CGGCGGCACCTTTCCTCTAATCCAGCAAAGTGATTCCCTGCACACCAGAGACAAGCAGAG 76

Qy    2882 taacaggatcagtgggtctaagtgtccgagacttaacgaaaatagtatttcagctgcaat 2941
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75 TAACAGGATCAGTGGGTCTAAGTGTCCGAGACTTAACGAAAATAGTATTTCAGCTGCAAT 16

Qy    2942 aa 2943
           ||
Db      15 AA 14


RESULT  98
AI033342/c
LOCUS        AI033342       482 bp    mRNA            EST       28-AUG-1998
DEFINITION   ox02d12.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA
             clone IMAGE:1655159 3', mRNA sequence.
ACCESSION    AI033342
VERSION      AI033342.1  GI:3254295
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 482)
  AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
  JOURNAL    Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Tel: (301) 496-1550
             Email: Robert_Strausberg@nih.gov
             This clone is available royalty-free through LLNL ; contact the
             IMAGE Consortium (info@image.llnl.gov) for further information.
             Insert Length: 749   Std Error: 0.00
             Seq primer: -40m13 fwd. ET from Amersham
             High quality sequence stop: 455.
FEATURES             Location/Qualifiers
```

```
     source          1. .482
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1655159"
                     /clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     This is a subtracted version of the original Soares fetal
                     liver spleen 1NFLS library.   1st strand cDNA was primed
                     with a Pac I - oligo(dT) primer [5'
                     AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       94 a    130 c    122 g    136 t
ORIGIN


  Query Match             6.1%;  Score 180;  DB 15;  Length 482;
  Best Local Similarity   100.0%;  Pred. No. 9.9e-84;
  Matches  180;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2502 ccacaggccaagaaggtcagagcccagtgaagatctgggagaccctgaactcagaaggct 2561
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    456 CCACAGGCCAAGAAGGTCAGAGCCCAGTGAAGATCTGGGAGACCCTGAACTCAGAAGGCT 397

Qy   2562 gtgtgtcttctgccccacgcacgcacccgtatctgccctccttgctggtagaagctgaag 2621
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    396 GTGTGTCTTCTGCCCCACGCACGCACCCGTATCTGCCCTCCTTGCTGGTAGAAGCTGAAG 337

Qy   2622 agcacggtcccccaggaggcagctcaggataggtggtatggagctgtgccgaggcttggg 2681
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    336 AGCACGGTCCCCCAGGAGGCAGCTCAGGATAGGTGGTATGGAGCTGTGCCGAGGCTTGGG 277


RESULT 99
BE619874/c
LOCUS       BE619874       577 bp    mRNA            EST       24-AUG-2000
DEFINITION  601473130T1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3876223 3',
            mRNA sequence.
ACCESSION   BE619874
VERSION     BE619874.1  GI:9890812
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 577)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM623  row: k  column: 08
            High quality sequence stop: 577.
FEATURES             Location/Qualifiers
     source          1. .577
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3876223"
                     /clone_lib="NIH_MGC_68"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
BASE COUNT       94 a    189 c    145 g    149 t
ORIGIN


  Query Match             6.1%;  Score 179;  DB 110;  Length 577;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-83;
  Matches  179;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   2594 ctgccctccttgctggtagaagctgaagagcacggtcccccaggaggcagctcaggatag 2653
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    323 CTGCCCTCCTTGCTGGTAGAAGCTGAAGAGCACGGTCCCCCAGGAGGCAGCTCAGGATAG 264

Qy   2654 gtggtatggagctgtgccgaggcttgggctcccacataagcactagtctatagatgcctc 2713
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    263 GTGGTATGGAGCTGTGCCGAGGCTTGGGCTCCCACATAAGCACTAGTCTATAGATGCCTC 204

Qy   2714 ttaggactggtgcctggcacagccgcgggccaggaggctgccacacggaagcaagcaga 2772
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    203 TTAGGACTGGTGCCTGGCACAGCCGCGGGCCAGGAGGCTGCCACACGGAAGCAAGCAGA 145


RESULT 100
T72963
LOCUS       T72963         376 bp    mRNA            EST       02-MAR-1995
DEFINITION  yc65b06.r1 Stratagene liver (#937224) Homo sapiens cDNA clone
            IMAGE:85523 5', mRNA sequence.
ACCESSION   T72963
VERSION     T72963.1  GI:689638
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 376)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 954
            High qality sequence stops: 335 Source: IMAGE Consortium, LLNL This
            clone is available royalty-free through LLNL ; contact the IMAGE
            Consortium (info@image.llnl.gov) for further information.
            Insert Length: 954   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 335.
FEATURES             Location/Qualifiers
     source          1. .376
                     /organism="Homo sapiens"
                     /db_xref="GDB:502580"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:85523"
                     /clone_lib="Stratagene liver (#937224)"
                     /sex="male"
                     /dev_stage="49 years old"
                     /lab_host="SOLR cells (kanamycin resistant)"
```

/note="Organ: liver; Vector: pBluescript SK; Site_1: EcoRI
; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo
dT. Hepatectomy from normal male caucasian. Average insert
size: 1.1 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5'
GAATTCGGCACGAG 3' ~3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTTT 3'"

ASE COUNT 82 a 94 c 121 g 63 t 16 others
RIGIN

Query Match 6.0%; Score 178; DB 145; Length 376;
Best Local Similarity 100.0%; Pred. No. 1.1e-82;
Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
y   2208 attatgctgaaccacttcagccagcgctatgccaaggtccccctcttcagccccaacttc 2267
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      37 ATTATGCTGAACCACTTCAGCCAGCGCTATGCCAAGGTCCCCCTCTTCAGCCCCAACTTC 96

y   2268 agcgagaaagtgggagttgcctttgaccacatgaaggtctgctttggagactttccaaca 2327
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      97 AGCGAGAAAGTGGGAGTTGCCTTTGACCACATGAAGGTCTGCTTTGGAGACTTTCCAACA 156

y   2328 atgcccaagctgattcccccactgaaagccctgtttgctggcgacatcgaggagatgg 2385
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     157 ATGCCCAAGCTGATTCCCCCACTGAAAGCCCTGTTTGCTGGCGACATCGAGGAGATGG 214
```

earch completed: February 18, 2001, 09:27:24
ob time: 14671 sec